January 2016

FINITE ELEMENT AND IMAGING APPROACHES TO ANALYZE MULTISCALE ELECTROTHERMAL PHENOMENA

Amir Koushyar Ziabari
Purdue University

Graduate School Form 30
(Revised 08/14)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Amirkoushyar Ziabari

Entitled: Finite Element and Imaging Approaches to Analyze Multiscale Electrothermal Phenomena

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

ALI SHAKOURI

CHARLES A. BOUMAN

GANESH SUBBARAYAN-SHASTRI

MARK S. LUNDSTROM

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

ALI SHAKOURI

Approved by Major Professor(s): _____

_____

Approved by: V. Balakrishnan 07/25/2016

Head of the Department Graduate Program Date

FINITE ELEMENT AND IMAGING APPROACHES TO ANALYZE MULTISCALE ELECTROTHERMAL PHENOMENA

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Amirkoushyar Ziabari

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2016

Purdue University

West Lafayette, Indiana

To My Wife, Maryam, for her indefinite kindness, patience, love, and support.

To My Parents, Sonia and Farrokh, who paved the way for me to achieve my goals.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my adviser Prof. Ali Shakouri for his constant support and encouragement. His positive attitude, flexibility, trust, and guidance helped me all along throughout my PhD and directed me to make the right decisions and advance in my research. I cannot be thankful enough for all he has done for me and hope to do something someday to make him proud of being my adviser.

I would like to express my sincere gratitude to my PhD committee, Prof. Mark Lundstrom, Prof. Charlie Bouman and Prof. Ganesh Subbarayan, for allotting the time to guide me whenever needed, and for their invaluable feedback throughout my PhD studies.

My QUEST groupmates and collaborators were not only supportive colleagues but also amazing friends without them I could not make this far. Kaz Yazawa, J.H. Bahk, J.H. Park, Bjorn Vermeersch, Xi Wang, Gilles Pernot, Kerry Maize, Yeerui Koh, Amr Mohammed, Yu Gong Wang, Alexander Shakouri, Xufeng Wang, Doosan Back, Harsha Eragamreddy, Hossein Pajouhi, Sajid Choudhurry and many other wonderful colleagues that cannot be listed in one paragraph.

The staff at Birck Nanotechnology Center at Purdue University are amazingly helpful and kind people and I really appreciate their service during my residence at Birck.

I want to thank my parents Sonia Masoumi and Farrokh Ziabari for their love and support in my entire life, and helping me finding and following this path. I would also like to thank

my siblings Kavian and Kamiar, and my extended family, my in-laws, Parviz, Gita, Moti, Mina, and Ali for their support and encouragement.

Last but not least, I wish to thank my wife Maryam Parsa. I have not been able to find a word to describe how grateful I am to your unconditional love and nonpareil devotion and support. Thank you for pushing me when I needed, thank you for providing a calm and friendly environment at home so I could follow my dreams, and thank you for standing by me patiently throughout this journey.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Appendix Figure

Appendix Figure Page

ABSTRACT

Ziabari, Amirkoushyar. Ph.D. Purdue University, August 2016. Finite Element and Imaging Approaches to Analyze Multiscale Electrothermal Phenomena. Major Professor: Ali Shakouri

Electrothermal effects are crucial in the design and optimization of electronic devices. Thermoreflectance (TR) imaging enables transient thermal characterization at submicron to centimeter scales. Typically, finite element methods (FEM) are used to calculate the temperature profile in devices and ICs with complex geometry. By comparing theory and experiment, important material parameters and device characteristics are extracted. In this work we combine TR and FEM with image blurring/reconstruction techniques to improve electrothermal characterization of micron and nanoscale devices.

We present an ultrafast yet highly accurate technique, based on image blurring and FEM, for thermal modeling of 3D ICs that include thermal vias. Modeling shows the impact of thermal vias placement on the reduction of maximum temperature in different layers of 3D IC. Next, we experimentally investigate high field non-equilibrium in electron gas and its impact on thermoelectric cooling. When there is a large temperature gradient or a high current density, thermoelectric coefficients can become non-linear. We provide the first detailed experimental study of non-linear Peltier coefficient in low-doped InGaAs

semiconductor. A hybrid analytical-numerical model is developed to extract current-dependent Peltier coefficient from TR thermal images of nonlinear InGaAs microrefrigerators at room and cryogenic temperatures. Finally, we combine FEM, image deconvolution techniques and TR to accurately reconstruct thermal images of devices with spatial resolution below visible light diffraction limit. We present full thermal distribution around submicron size heat sources on InGaAs thin films on InP substrate both in steady-state and in transient regime indicating emergence of non-diffusive heat transport.

1. INTRODUCTION

1.1. Introduction

When electrons flow through a conductor their collisions with atoms in the lattice lead to irreversible heat generation and in turn, temperature elevation. The average energy of electron gas also changes when carriers move from one conductor to another and this can lead to reversible Peltier heating and cooling. These interactions are the foundation to a class of phenomena, called electrothermal phenomena. Studying these effects has been an integral part of research in a diverse spectrum of technologies such as electronic devices and integrated circuits, thermoelectric energy conversion, phase change memory, and microelectromechanical systems (MEMS), to improve efficiency, performance, and reliability.

According to Moore's law, as the feature size of transistors shrinks, the number of transistors in integrated circuits (IC) will double approximately every two years [1]. This has resulted in a large on-chip power density, evolving hot-spots, and temperature non-uniformity across the chip. Such localized high temperatures may cause reliability phenomena such as Negative/Positive Bias Temperature Instability (NBTI/PBTI) and Time Dependent Dielectric Breakdown (TDDB), and hence accelerate the degradation of transistors. Additionally, the leakage power will increase. Leakage power is exponentially

dependent on the temperature and can degrade the performance or even cause complete circuit failure.

Three dimensional vertically integrated circuits (3D ICs) offer a promising solution to continue the CMOS performance trend in the next decade. Stacking multiple device layers in 3D ICs enables higher transistor densities and shorter interconnect lengths than planar (2D) ICs [2]. Due to an increase in the die effective area and also presence of low thermal conductivity interlayers, the power density and the thermal resistance from die to package will increase, and thus, thermal issues are expected to be a key limitation in high performance 3D ICs [3], [4]. Such electrothermal effects impose restrictions on IC designers to maintain the reliability and to deal with power management issues, requiring close collaboration between chip and package designers and electrothermal co-optimization to achieve desirable circuit performance.

Thermoelectric (TE) materials have been employed for energy conversion applications for many years. In order to increase the efficiency of TE materials, one has to enhance their figure of merit ($ZT = \frac{S^2\sigma}{\kappa}T$). Enhancing ZT implies increasing electrical conductivity (σ) and Seebeck coefficient (S), which are inversely related, and decreasing thermal conductivity (κ). Due to their poor efficiency, TE energy conversion devices had limited niche applications. Since the early nineties, the interest towards thermoelectrics has increased significantly after it was theoretically predicated that semiconductor nanostructuring could significantly improve ZT [5]–[7]. By introducing complex structures and nanocomposites one can engineer material properties at nanoscale. Lattice thermal conductivity will be reduced by scattering phonons at the boundaries using superlattice

materials or with embedding nanoparticles distributed with periods/spacing smaller than the phonon mean free path, while electronic contribution to thermal conductivity remains unchanged; hence, the total thermal conductivity will be reduced with minimal decrease in electrical conductivity [8]–[12]. Further, through electron filtering and modification of the density of states, Seebeck and electrical conductivity can be decoupled in these nanostructures, leading to an enhancement in Seebeck effect and improved figure of merit (ZT) [12]–[14]. In addition to optimizing the material properties, integration of the TE materials into reliable, and cost-effective thermoelectric modules for application such as localized cooling of hot spots [15]–[18] and waste heat recovery [15], [19] demands a thorough optimization of device geometry and design [20]–[22]. Electrothermal effects play a crucial role in optimizing the performance of the TE materials, as well as in designing a mechanically robust TE module when taking thermally-related reliability concerns into consideration.

Numerical and analytical modeling are required to investigate electrothermal effects in semiconductor materials and devices. Due to nonlinear interactions and close correlation between power dissipation, temperature, performance and reliability, modeling is an essential element in design optimization and accurate estimation of each of these parameters. For instance, at the early stages of the IC design, electrothermal modeling plays a significant role in thermal aware design (floorplanning), placement and routing of functional blocks [3], [23], [24]. Conventionally, finite element (FE) numerical method is used to find a near exact solution to boundary value problem for the heat diffusion equation. This requires the entire domain (the device structure as well as heat spreader) to be

discretized to smaller subdomain (elements). The variational method is then used to minimize the total potential energy functional, which is a summation of the energy contribution of each individual element [25].

The accuracy of finite element method comes at the cost of increasing the mesh density of the structures. The denser the mesh, the more accurate the results and the longer the computation time. For the complex structures in planar and 3D ICs, it is impractical to use FEM for thermal aware floorplanning, and placement and routing of functional blocks in ICs. Therefore, it is necessary to develop techniques that are capable of analyzing and predicting the thermal behavior of an IC with fast computation and high accuracy. In Chapter 2, we present a fast convolution-based technique, deemed Power Blurring (PB), for static and dynamic thermal modeling of planar (2D) and 3D ICs. In particular, we emphasize thermal modeling of 3D ICs with thermal vias. The PB method will be compared against FEM in terms of accuracy and speed of computation.

In thermoelectrics (TE) applications, electrothermal modeling is critically important for material and design optimization to maximize the efficiency, and to avoid reliability issues due to the thermal stress induced fracture or delamination. Thermoelectric generators (TEGs) are emerging as a possible solution for high temperature energy conversion applications and waste heat recovery systems. The key challenges are improving the efficiency of thermoelectric power generator module (TEM) and its material cost in large scale production. A system optimization for TE waste heat recovery system and minimization of the TEM cost is presented in Ref. [20]. The closed form analytical solution reveals that the optimum solution for the maximum output power can be obtained by

thermal impedance matching with heat source and the heat sink (hot and cold reservoirs) as well as electrical impedance matching with the load. Upon finding the optimum solution, cost-performance analysis is conducted to find the minimum cost design at a given system efficiency. This optimization elucidates that the fractional area coverage of the thermoelectric (TE) leg, called fill factor (FF), plays a significant role in minimizing the mass of the TE material used in thermoelectric waste heat recovery systems. It is shown that improving the figure-of-merit (ZT) along with decreasing the fill factor would further reduce the total cost [20]. Because the maximum power output from a thermoelectric system is proportional to square of temperature difference between the hot and cold reservoirs [22], employing a thermoelectric generator with optimum design in high temperature applications and with large temperature difference, such as on top of a steam turbine cycle, will be an economical approach to increase energy production [21]. However, both reduced fill factor and higher temperature range imply a larger impact on reliability of TEMs. In Chapter 3, we employ analytical modeling and FEM to investigate the impact of lower fill factor and high temperature on thermally induced interfacial shearing stress, one of the factors responsible for mechanical stability of TE module. Additionally, we utilize FEM to study the impact of variations in the geometry, boundary conditions, and the coefficient of thermal expansion (CTE) of the materials on the maximum shearing stress, focusing on the fill factor, in a thermoelectric power generator module (TEM) designed for high temperature applications.

Akin to theoretical modeling, experimental characterization techniques are indispensable in analysis and understanding of the electrothermal effects in semiconductor materials and

devices. One has to perform thorough measurements and analysis to extract material properties and in turn estimate the performance of semiconductor devices. Moreover, electrothermal phenomena such as non-uniform current and heat distribution and consequently localized high temperature or hot spots limit the operational temperatures of semiconductor devices and may degrade their performance [26]. These effects can only be revealed by techniques that are able to measure the full spatial distribution of temperature profiles in devices.

Thermoreflectance (TR) microscopy is a noninvasive optical technique that is suitable for the 2D mapping of temperature field in active semiconductor devices [27]. TR detects a small change in the sample surface reflectivity due to a change in the temperature. The reflectivity (R) of a material is related to its temperature change (ΔT) through the thermoreflectance coefficient (C_{TR}), by equation (1).

$$\frac{\Delta R}{R} = \left(\frac{1}{R}\frac{\partial R}{\partial T}\right)\Delta T = C_{TR}\,\Delta T \tag{1.1}$$

The magnitude of the thermoreflectance coefficient (C_{TR}) is usually on the order of 10^{-2} - 10^{-5} (1/K) depending on the materials measured. This coefficient is influenced by sample composition, wavelength of probing light, and numerical aperture of the imaging system [28]. TR microscopy can be employed for both quasi-static and transient 2D temperature profiles of devices. This method is sensitive to mK temperature variations with sub-nanosecond temporal resolution, and can render images with submicron spatial resolution [27], [29]–[33]. In Chapter 4, we exploit TR to experimentally investigate high field non-equilibrium in electron gas and its impact on thermoelectric cooling.

Thermoelectric cooling occurs due to the Peltier effect that is the change in the average energy of electrons when current flows through an interface between two dissimilar conductors or semiconductors. When current traverses a junction between two conductors A and B, where in A the average transport energy of electrons is lower than B, electrons will absorb heat from the lattice at the junction to compensate for their lower energy and therefore cool down the junction. This is called Peltier cooling. The amount of Peltier heat carried at the junction is proportional to the applied current according to the Peltier coefficient Π. The Peltier coefficient is independent of the electrical current and is related to the Seebeck coefficient by the Kelvin relation, $\Pi = S \times T$. Using Monte-Carlo simulations, it is demonstrated that under strong electric field, electronic temperature starts to exceed the lattice temperature. This non-equilibrium electron heating leads to nonlinearity in the Peltier coefficient. It is predicted that nonlinearity is proportional to effective mass, square of current density and inversely proportional to square of the carrier concentration of the semiconductor [34]. Additionally, due to weaker electron-phonon coupling, the nonlinearity will be more prominent at lower temperatures. It is predicted that nonlinearity can be utilized to enhance thermoelectric cooling at large current densities [34].

In an effort to measure the anticipated nonlinear Peltier coefficient in low doped semiconductors and the corresponding cooling, a set of thin film InGaAs microrefrigerators is designed and fabricated. The temperature change on the top of these devices is obtained using thermoreflectance thermal imaging at different current densities. Exploiting TR thermal imaging results and the four-point probe electrical characterization, at room and

cryogenic temperatures (30-70K), along with a hybrid analytical-numerical model, we extract the reversible non-linear Peltier coefficient at different current densities and separate it from irreversible non-linear Joule heating in the device. In chapter 4, we describe the methodology and present the extracted Peltier coefficient. We also discuss how this will affect the amount of the Peltier cooling/heating and in turn the performance of microrefrigerator devices at room and cryogenic temperatures.

Exploring electrothermal effects at nanoscale is critical for many current and future applications of nanostructures. Reducing the size of devices to the nanoscale and engineering the nanoscale material properties is used to enhance the performance of the ICs and TE materials and devices. At smaller dimensions comparable to electron or phonon mean free paths in the materials, the charge and heat transport cannot be treated based on classical diffusive model and their non-diffusive behavior needs to be considered. It has been predicted and shown that Fourier diffusion equation fails to explain the thermal behavior [35]–[38]. Nonlocal and nonequilibrium effects due to phonons carrying heat ballistically away from the heat source are the main cause of the departure from the Fourier law. In such a non-equilibrium condition at nanoscale, where assumptions of continuum and second law of thermodynamic may not hold, questions emerge concerning the definition of temperature and its measurements. To answer the question “whether we are able to define temperature at the nanometer scale?”, Hartman et. al. analyzed a harmonic chain of quantum particles with next neighbor interactions and found minimum length scales at which the local temperature can be defined at different ambient temperatures [39], [40]. About 0.1m at 1K and 100nm at 680K are the minimum length scale predicted for 1D

silicon atoms. Although these numbers seem to be oddly large, this may only be due to the 1D nature of their calculations and ignoring phonon scattering events. Nevertheless, their approach captures the essential physics and can be a path toward finding more accurate estimation models. In the cases where local equilibrium assumption is no longer valid, then one needs to modify the equilibrium concepts such as entropy to be able to define temperature in non-equilibrium situations. Such approaches, e.g. Extended Irreversible Thermodynamics (EIT), are discussed in the review done by Vazquez et. al. [41] and more recently in [42]. Experimental investigations using laser based Time Domain Thermoreflectance (TDTR) [43]–[47] and Frequency Domain Thermoreflectance (FDTR) [48] techniques, using Coherent Soft X-ray beams [49], as well as using $2\omega/3\omega$ electrical measurement [50], have probed the non-diffusive thermal transport in semiconductor materials and alloys. In these experiments, a heat source such as a laser source or Joule heating due to current flow through a metal heater are used to heat up the substance under study. Because of the fast electron-electron interactions in the metal transducer, the sensor temperature can often be defined unambiguously. An important question is to relate temperature variation of the sensor to the "energy" transport in the crystal and if a unique temperature could be defined for various phonon modes. When the size of the heat source is large, phonons emitted from the heat source go through a random walk, scatter, change the equilibrium entropy of the system and therefore temperature change in a classical sense can be defined. Reducing the size of the heat source or increasing the modulation frequency result in quasiballistic phonon transport in which some of the phonons maybe in non-equilibrium (e.g. ballistic transport of long mfp phonons in Si [45]) while others scatter

and dissipate energy which translate to change in the non-equilibrium entropy of the system and therefore translate into the temperature. In an alternative explanation, long jumps are intermingled with short random walks in a Lévy flight picture which does not require introducing different temperature for ballistic and diffusive model (see [46]).

To truly understand the physics behind the heat transport at nanoscale, and to investigate various theories described in aforementioned references, versatile thermal characterization technique with high thermal, spatial and temporal resolution is necessary. TR spectroscopy is proven useful for 2D mapping of thermal effects at nanoscale and qualitative elucidation of physical phenomenon at small scales [51]–[53]. TR is employed to create 2D temperature map of electrically biased single crystalline VO_2 nanobeams, and to observe strongly localized alternating Peltier heating and cooling as well as Joule heating at the Metal-Insulator domain boundaries [51]. Furthermore, the technique is used to observe super-Joule self-heating at the transport bottlenecks in networks of silver nanowires and silver nanowire/single layer graphene hybrid films [52]. Figure 1.1 shows the optical images of a GaN HEMT device with 0.1 micron wide gate metal under 20x and 100x objective lens magnification with overlay of thermal intensity obtained by TR microscopy when one of the two parallel gates is biased [54]. Although, illumination with visible wavelengths enables TR to have a high spatial resolution compared to methods such as IR thermal imaging, the optical diffraction still limits accurate quantitative temperature profiles in devices with feature sizes below Rayleigh resolution criterion [55]. In Chapter 5, we first describe why TR thermal imaging is capable of acquiring the thermal images below diffraction limit despite the fact that the optical images are obscured at those scales.

Then, we combine FEM, image processing, and TR thermal imaging to accurately reconstruct thermal images of devices at scales below diffraction limit. We present a maximum *a posteriori* (MAP) framework [56], using Markov Random Field priors [57], to reconstruct true thermal images from the diffraction-limited ones. Reconstruction is performed both on *numerically-designed experiments* and true thermal images. Independent measurements to verify the results are performed. In Chapter 6, we present TR imaging results to study nanoscale heat transport in heater lines fabricated on InGaAs semiconductor alloys, and to examine superdiffusive heat transport model presented in Refs [46], [58]. Room and cryogenic temperature measurements were performed at both steady-state and transient regimes. Signatures of non-diffusive heat transport are observed and analyzed in these measurements. In Chapter 7, we provide some proposals for future prospect and continuation of the research presented in this dissertation.

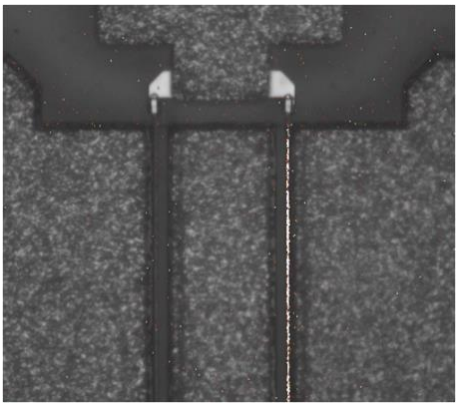
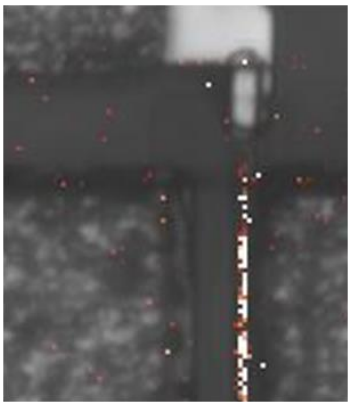
(a)  (b) 

Figure 1.1. Optical image and thermal intensity overlay of a GaN HEMT device. The magnifications are (a) 20x and (b) 100x. The gate length is 35 μm. The top surface gate contact is approximately 150 nm wide and the gaps to the drain and source contacts are approximately 400nm and 200 nm, respectively.

2. POWER BLURRING IN 3D ICs INCLUDING THERMAL VIAS

2.1. Introduction

In recent years, the scaling of supply voltage has departed from the ideal scaling predicted by Moore [59] and Dennard et al. [60]. The threshold voltage (V_{th}) has stopped scaling. This in turn has stopped the scaling of supply voltage (V_{dd}) to maintain circuit performance. This results in higher power densities, which promotes temperature as one of the primary design parameters as the scaling advances to smaller feature sizes. In addition, non-uniform activities in an IC chip yield non-uniform surface temperature distributions, since localized heating occurs much faster than chip-wide heating [61]. Thus the temperature of certain regions can become much higher than that in the neighboring areas. These higher temperature regions are called hot spots. Hot spots and spatial temperature gradients in VLSI ICs have become critical issues due to their limiting effects on both the performance and the reliability of packaged IC chips[62].

Moreover, the increasing leakage power and its exponential dependence on the temperature require more attention to thermal-aware simulations and optimizations. Hence, precise estimation of temperature distribution is essential for accurate analysis of performance, reliability, and power management. Generally, thermal simulations and design optimizations are done under steady-state worst-case conditions due to the high computational cost, causing reliance on the use of conservative margins in thermal designs.

However, the temperature non- uniformity evolves over time and thus hot spots are of spatio- temporal nature. The transient temperature spike or localized heating also can cause timing errors, non-uniform current flow or reliability failures. As the thermal budget becomes increasingly tight, the worst-case approach becomes too costly and ineffective. It is also known that the simulated worst-case peak power and its corresponding peak temperature are rarely observed [26], [61].

Even with the state-of-the-art tools, the chip-level transient thermal simulation with a realistic package configuration is too costly for physical design optimization or performance verification in the packaged environment. In addition, in the early stages of chip design, i.e. architectural specification stage, specific package information and thermal boundary conditions may not be available. At these stages, designers rely on simulation to consider the trade-offs. However, slow simulation limits the scope of analysis. For these reasons, a fast transient thermal analysis method is highly desired [63].

Three dimensional vertically integrated ICs (3D ICs) offer a promising solution to continue the CMOS performance trend in the next decade. Stacking multiple device layers in 3D ICs enables realizing higher transistor densities and shorter interconnect lengths than two-dimensional (2D) ICs [2]. Due to increase in the die effective area and, in turn, the power density and thermal resistance, thermal issues are expected to be a key limitation in high performance 3D ICs. This prompted the IC designers to incorporate electrical vias into the 3D ICs to lower the interconnect length between device layers. Since vias are good thermal conductors as well, placing “thermal vias” between device layers and heat sink is effective to lower the temperature [64]. Figure 2.1 shows a typical 3D stacking of multiple active

layers inside a single package. “Thermal vias” can be inserted to remove heat from bottom device layers to the heat sink while “power/ground vias” are used to carry power supply between active layers [64].

As a result of thermal challenges such as temperature non-uniformity, thermal-aware design has become an essential aspect of 3D IC design. Accurate estimation of the temperature profile in each active layer is very important. For example, the need for fast thermal analysis is indisputable in placement and routing of active blocks and/or thermal vias in 3D ICs [65].

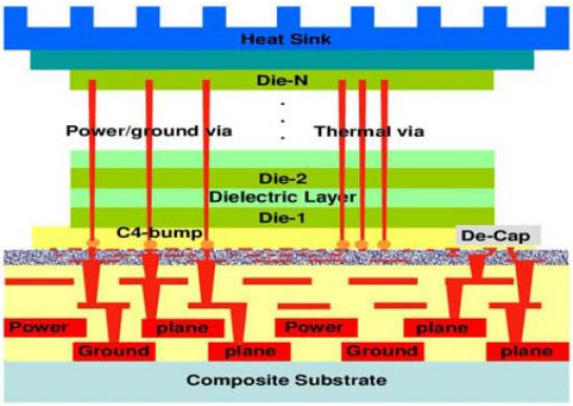


Figure 2.1. A typical 3D stacking with thermal vias [64].

In the reminder of this chapter, we present a fast, yet accurate steady-state and transient temperature computation method suitable for VLSI planar (2D) and 3D ICs in packages. In order to validate the new method, called Power Blurring (PB), simulation results are compared to those of a Finite Element Analysis obtained by ANSYS [66]. The main emphasis is on 3D ICs including thermal vias. In section 2.2 we describe the Power

Blurring methodology for fast static and transient thermal modeling in ICs. In Section 2.3, we present the algorithm for Power Blurring in 3D IC including thermal vias followed by case study results and discussion.

2.2.Power Blurring (PB) Method

To obtain the temperature profile in a chip, the heat diffusion equation shown in Equation (2.1) has to be solved with appropriate boundary conditions

$$k\frac{\partial^2 T}{\partial x^2}+k\frac{\partial^2 T}{\partial y^2}+k\frac{\partial^2 T}{\partial z^2}+q^*=\rho c_p\frac{\partial T}{\partial t} \tag{2.1}$$

where k is the thermal conductivity (W/m·K), ρ is the density (kg/m^3), and c_p is the specific heat (J/kg·K), q^* is the heat generation per unit volume (W/m^3), and T is the temperature of the location (x, y, z) at time t. In 3D IC thermal analysis, the heat diffusion equation has been conventionally handled by grid-based methods, such as Finite Difference Method (FDM) or Finite Element Method (FEM), which generate 3D volumetric meshes of solid structure. The accuracy of these FDM and FEM methods comes at the price of long execution times, and exhaustive CPU and memory usage. A matrix convolution technique, called *Power Blurring (PB),* has been developed to expedite the computation of temperature distribution in IC chips. This PB method has its theoretical basis on the Green's function method [67], [68]. Implementation is similar to image blurring used for image processing.

Equations (2.2) and (2.3) govern the Green's function method and its equivalent implementation in PB, respectively. In (2.2), $G(\rho,\rho')$ is the Green's function, and V' is the volume over which the heat, $q(x',y',z')$, is generated.

$$\Delta T = \iiint G(\rho, \rho')q(x', y', z')dV' \tag{2.2}$$

Thus the Green's function (response to a delta-function excitation) is a building block to construct the solution.

In image processing, an image is blurred by a convolution with a filter mask. The filter mask is a matrix whose elements define a process in which the modification (i.e. blurring) of an image occurs. For instance, an image, *f*, is convoluted with a filter mask, *w*, to produce a blurred image *g* by the Equation (2.3).

$$g(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x + s, y + t) \tag{2.3}$$

where *a=(m-1)/2* and *b=(n-1)/2* for a $m \times n$ mask [56].

In thermal analysis, if we think of the power map as an image, *f*, the resulting temperature profile can be regarded as a blurred image, *g*, where the filter mask, *w*, corresponds to the thermal mask.

Table 2.1 summarizes the analogy between the image blurring and the PB methods.

TABLE 2.1. ANALOGY BETWEEN IMAGE PROCESSING AND POWER BLURRING.

Image Processing	Power Blurring
Image (*f*)	Power Map
Filter Mask (*w*)	Impulse Response
Blurred Image (*g*)	Temperature Profile

As mentioned earlier, thermal mask is the impulse response of the system in space domain. According to the Green's function method, we can build a solution to a partial differential equation with an arbitrary driving function once we have a solution corresponding to an impulse (a point source). In thermal analysis of IC chips, temperature distribution is the physical quantity of interest, and heat (i.e. power consumption in ICs) is the driving

function. Thus, the thermal mask conforms to a steady-state temperature distribution induced by a point heat source which is applied to the center of the Si die. In practice, the die area is discretized into grids and an approximate delta function simulating a point heat source is applied to the center element of the grid. Subsequently, the difference between the resulting temperature and ambient temperature is normalized with the amount of the input power. Although the thermal mask can be obtained in analytical form for a simple 1D geometry, measurement or 3D finite element analysis (FEA) simulator is needed for a realistic structure shown in Figure 2.2, where 3D heat spreading in the package is important. Figure 2.3 shows the thermal mask for a given package model. Its unit is that of thermal resistance (°C/W), and hence the thermal mask generates a temperature profile when it is convoluted with a power distribution map (W).

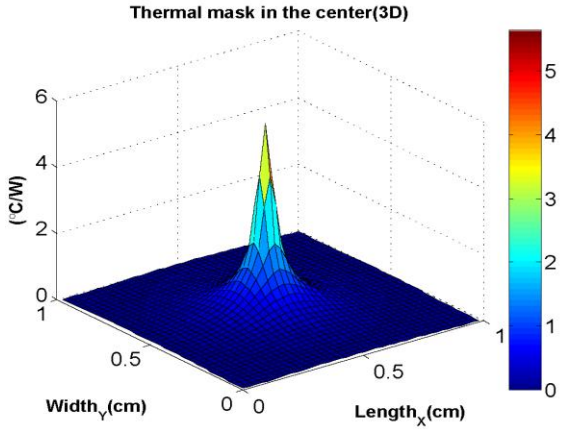


Figure 2.2. The thermal mask; the surface of the Si die is discretized into 40×40 grids.

The shape of the thermal mask depends on the location of the point heat source. Method of images is developed to avoid necessity of obtaining thermal mask at different locations of

the chip. The principle of the method of image for power blurring is similar to Method of Image in electromagnetism [69]. Consider the case of a heat source located at a distance, d, from an adiabatic surface as shown in Figure 2.3a. No heat transfer can occur at the adiabatic boundary. This means the net outgoing heat flux at x = 0 and x = L is zero. If we replace the adiabatic boundary with an image source, the adiabatic boundary condition is satisfied. The problem then becomes more manageable with this approach. One need to convolve the extended power map with the thermal mask, obtained at the center, to calculate the temperature profile rather than using different thermal masks at edges, center and corner of the chip.

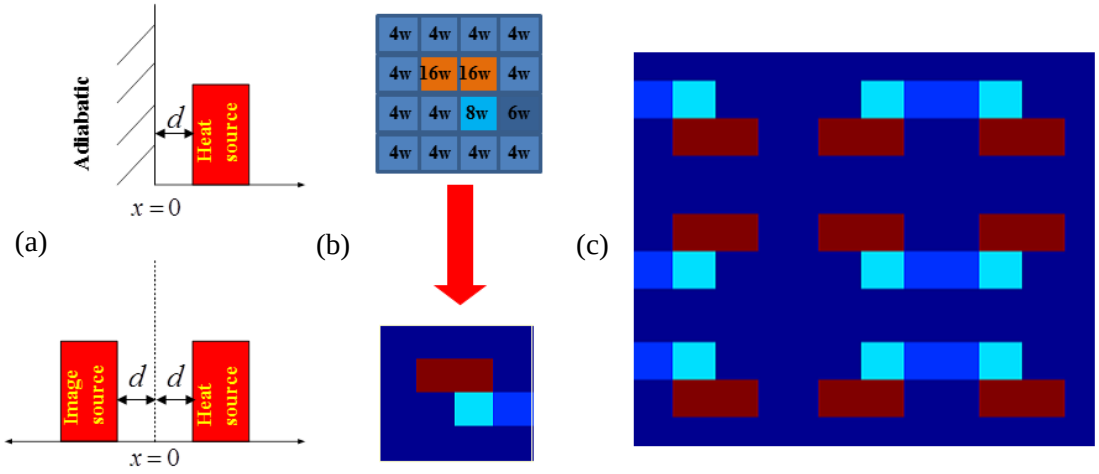


Figure 2.3. Method of Image. (a) Method of Image. (b) An arbitrary power map on Si die. (c) Power map with its eight nearest neighbor mirror images.

The extended power map can be calculated from equation (2.4).

$$P_{image} = [P(x,y) + P(x,-y) + P(-x,y) + P(x,2W-y) + P(-x,-y) + P(-x,2W-y) + P\,(2L-x,y) + P(2L-x,-y) + P(2L-x,2W-y)] \quad (2.4)$$

Here, P_{image} is the extended power map, P is the power map, and W and L are the width and the length of the power map, respectively. P_{image} extends from (-W, -L) to (2W, 2L) which is 9 times larger than P. For final calculation of temperature profile two third of this matrix is only needed.

The direct convolution result is given by

$$T_{DCMI}(x,y) = h(x,y) * [P_{image}(-\frac{L}{2}:\frac{3L}{2}, -\frac{W}{2}:\frac{3W}{2})] \quad (2.5)$$

where, T is the thermal profile, and h is the thermal mask shown in Figure 2.2. This is called Direct Convolution with Method of Image (DCMI). DCMI does not handle the 3D heat spreading at the boundary properly. Another error reduction step is added, which is called intrinsic error reduction, in which the results obtained by DCMI will be normalized by the intrinsic error between FEM and DCMI response to a uniform heat source [70]. Intrinsic error can be calculated from equation (2.6), in which $T_{DCMI\_Uniform}$ and $T_{ANSYS\_Uniform}$ are responses of DCMI method and ANSYS FEM to uniform heat source on top of the die.

$$E_r = (T_{DCMI\_Uniform} - T_{ANSYS\_Uniform})/T_{ANSYS\_Uniform} \quad (2.6)$$

The final temperature for an arbitrary thermal map can be calculated from equation (2.7).

$$T_{final} = \frac{T_{DCMI}}{1+E_r} + T_{ambient} \quad (2.7)$$

Where $T_{ambient}$ is the ambient temperature around the package.

It should be pointed out that although calculating thermal mask and intrinsic error require FEM, these calculations need to be done once offline for the package and used in a

convolution algorithm for thermal management and thermal aware floorplanning in ICs with arbitrary power dissipation profile.

A case study for temperature profile calculation using power blurring in an AMD Bobcat mobile processor model is shown in Figure 2.4. Figures 2.4a, and 2.4b, show the floorplan and the 2D power dissipation map for this processor. Figure 2.4c, shows the temperature profile obtained by power blurring. In Figure 2.4d, we compared power blurring results along vertical cross section with ANSYS and HotSpot thermal simulator [71]. HotSpot is a widely used thermal simulator in computer architecture community, and similar to other architectural level simulators, is designed to calculate temperature profiles which are accurate for the experiments at the architecture level (block sizes in tens of micro to millimeter range), and still fast enough to allow for the simulation of long dynamic temperature traces on the order of seconds. HotSpot is based on an equivalent circuit of thermal resistances and capacitances that correspond to the microarchitecture blocks. The essential aspects of the thermal package are also taken into account. In Table 2.2. a comparison in terms of accuracy and speed between PB, and HotSpot is carried out.

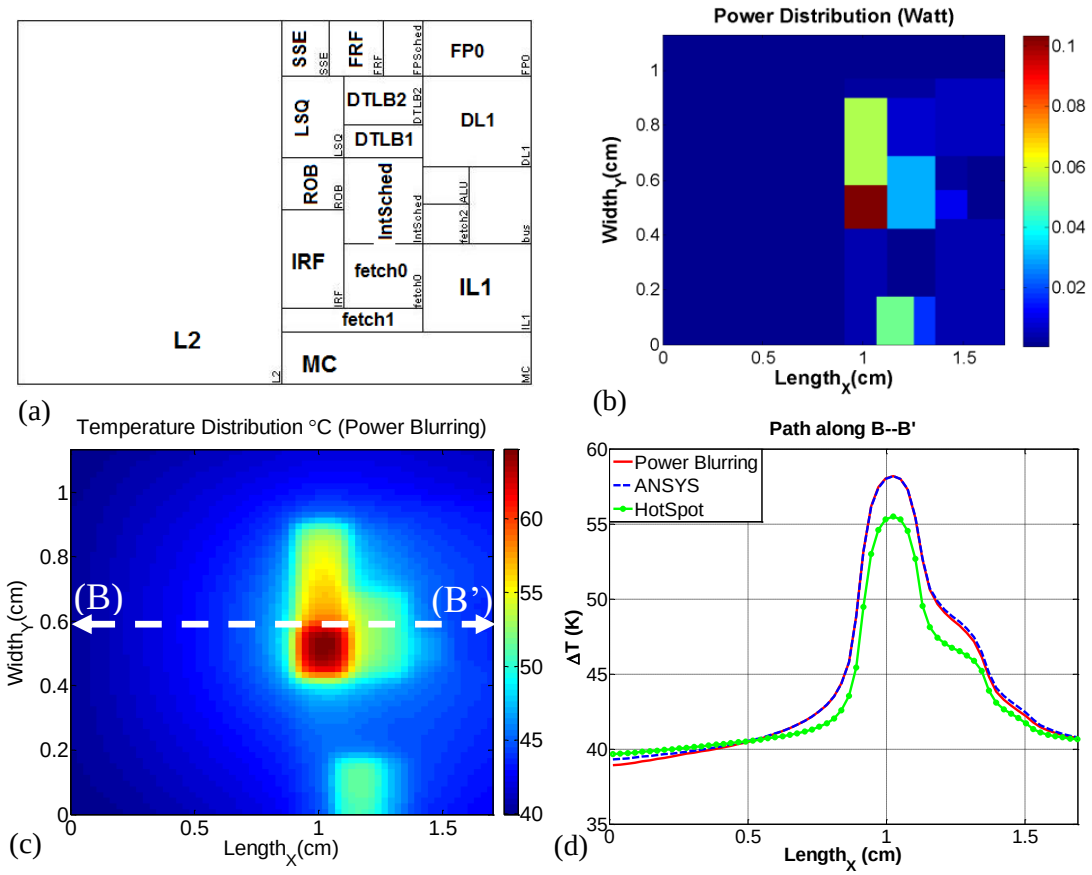


Figure 2.4. The floorplan and power map for steady-state case study. (a) The floorplan of the device; (b) A typical mobile processor model power map; (c) Temperature profile obtained by Power Blurring (PB); (d) Comparison between PB, ANSYS [66] and HotSpot [71] along B-B'.

Table 2.2. Comparison between HotSpot, and Power Blurring.

	ANSYS	HotSpot	PB
Computation Time	56s	0.11s	0.041s
Err. in hot-spot	-	12.9%	0.14%
Avg. Err	-	6.5%	2.5%
Abs. Err. range	-	0-4.2 ˚C	0- 0.56˚C

An adaptive power blurring technique is developed in [72] which can solve non-linear problems, e.g. when the thermal conductivity of the silicon is modified based on the local

temperature of the chip, using two or three iterations. Excellent agreements with self-consistent finite element simulations have been obtained.

PB further extended so that it can be applied for transient thermal simulations [70]. The difference is that the time evolution of the thermal mask resulting from spatiotemporal impulse is employed for the transient simulation. To obtain an impulse response in the time domain, a delta function needs to be applied to the center of the die. In practice, a point heat source is applied for a very short time period (approximate delta function), and the corresponding thermal response is recorded at each time step, which is shorter than the width of the approximate delta function. The width of the delta function is determined by the desired level of temporal resolution. The resulting thermal responses are normalized with respect to the amount of applied power. The series of thermal masks acquired at the end of this procedure constitutes a transient thermal mask (i.e. time evolution of the thermal mask). Once the transient thermal mask is prepared, the transient temperature profile is obtained by means of the superposition principle. A schematic overview of the transient thermal simulation process is presented in Figure 2.5. Three orders of magnitude and 28 times speed up compared to FEM and HotSpot simulator are shown, respectively, for a sample transient power dissipation map [70], [73].

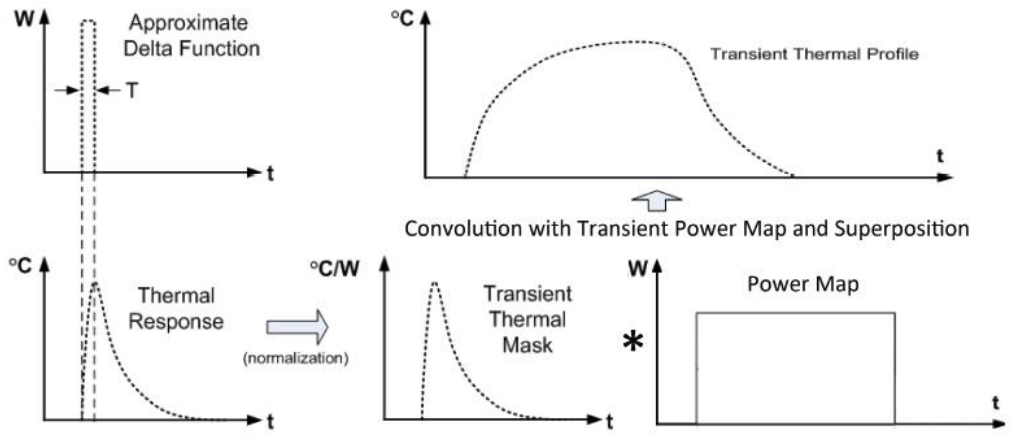


Figure 2.5. Schematic overview of power blurring transient thermal simulation.

2.3. Power Blurring (PB) in 3D ICs including thermal vias

We extended PB for thermal modeling in 3D ICs. In this case the temperature profile in each active layer is not only due to heat dissipated in that layer but also due to its neighboring active layers. Therefore, the coupling between neighboring layers must be considered. This is shown in Figure 2.6. In order to take this coupling into account, for e.g. a two layers chip, two thermal masks will be needed for each active layer. The first one is the temperature rise in the layer as a result a applying a point heat source in it and the second one is the temperature rise in its neighboring layer as a result of the same point heat source.

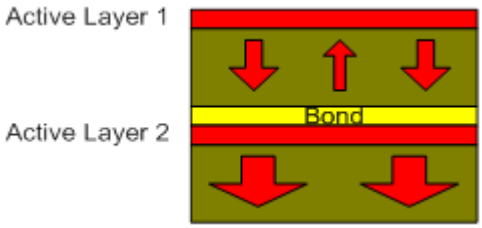


Figure 2.6. A Schematic diagram of heat flow in a two-layer 3D IC chip. Heat sink (not drawn) is on the bottom surface [74].

After obtaining thermal masks for each die, in a two-layer 3D IC, the following equations are used to calculate each layer's temperature profile:

$$T_1 = Mask_{11} * P_1 + Mask_{12} * P_2 \ \ and \ T_2 = Mask_{21} * P_1 + Mask_{22} * P_2 \qquad (2.8)$$

where P_1 and P_2 represent the power maps of layers 1 and 2, respectively.

For 3D ICs when there are thermal vias, the temperature distribution profile is significantly affected since a uniform silicon chip is not dominating heat spreading. One needs to calculate heat spreading in both silicon chip and the thermal via regions. We introduce two thermal masks: *Si thermal mask* and *via thermal mask*. Instead of convolving the entire power maps with the Si thermal masks, we need to convolve each element of power map with the appropriate thermal mask.

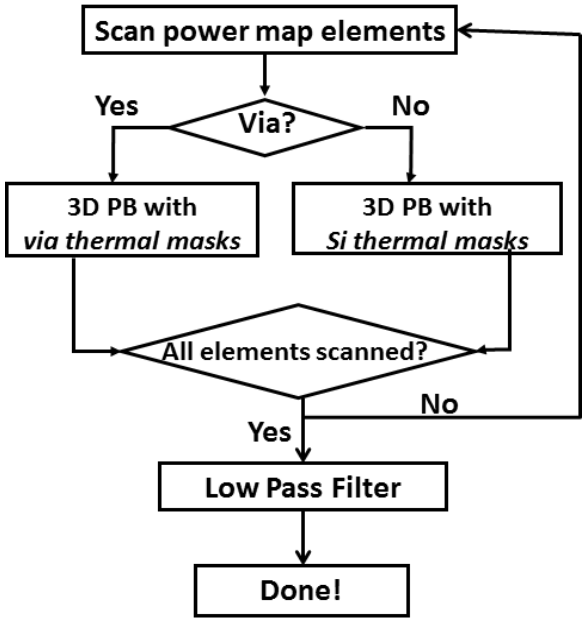


Figure 2.7. Algorithm for PB in 3D ICs with thermal vias.

The algorithm for PB in 3D ICs with thermal vias is shown in Figure 2.7. As it can be seen in this new method we should scan all the meshed elements and calculate the temperature profile at each element based on material properties of that element. Each meshed element in the power dissipation profile will be convolved with the appropriate thermal masks, i.e. Si thermal mask or via thermal mask, to calculate the temperature in that element. After temperature values over the entire chip are obtained then a low pass filter is used to smooth out the temperature profile at the edges of Si and via for each layer (this avoids using an additional thermal mask at the boundary between silicon and thermal via).

The significance of this method over the conventional FEA is in that, once the thermal masks for a geometry are calculated then for any power dissipation profile in each layer, the temperature profile can quickly be obtained and there is no need to re-mesh the entire

geometry. Also, comparing the method to 3D power blurring, we are now able to incorporate any non-uniformity in the chip which can be due to thermal vias, or even chip with different material properties in different regions. In the next section the results obtained for a case study and compared with ANSYS.

2.4. Case studies results and discussion

In the following case studies, we will follow the idea proposed in [65], in which the thermal vias are placed in designated regions called thermal via regions, and their density is adjusted with minimal perturbation on routing. The thermal conductivity of each of the regions is determined by the density of thermal vias in that region.

Performing full-chip thermal analysis at the granularity of individual thermal vias is unreasonable, because it will lead to very large matrix sizes in both FEA and Power Blurring. Therefore, it is necessary to find a way to alleviate computation complexity. In Appendix A.1., we illustrated that instead of modeling individual vias we can model the via region with the weighted average value of thermal conductivity of via material and the oxide.

Figure 2.8 shows the full chip thermal model. Material parameters and dimensions are summarized in Table 2.3. The configuration consists of two Si ICs with surface areas of 1cm×1cm and a Cu heat sink with a heat spreading layer. Thermal interface material (TIM) reduces the effect of surface roughness on the contact area, and hence improves heat conduction at the interface. Assuming flip chip package, most of the heat is considered to flow through the bottom surface of the heat sink, and other minor heat transfer paths are

neglected. Heat transfer coefficient of 0.15W/cm^2-K is employed for the bottom surface of the copper heat sink and adiabatic boundary condition is imposed on other surfaces. 3D IC technology is an emerging technology and bonding process technologies are still being developed. In this study, we assumed bonding approach of adhesive polymer BCB (Benzocyclobutene) and "face-to-back" bonding. 10% of the die area is dedicated to 16 thermal via regions.

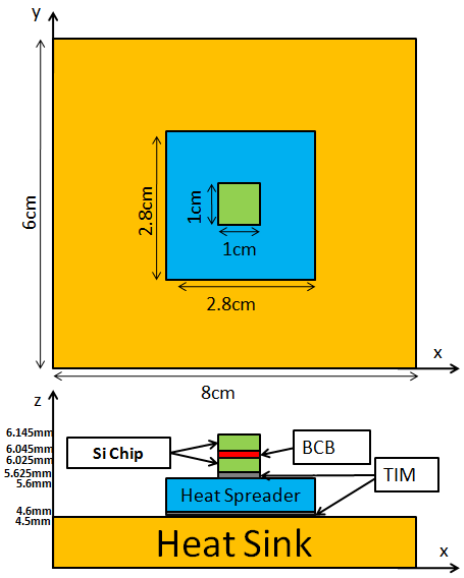


Figure 2.8. Schematic of a packaged 3D IC chip, where the heat spreader, the heat sink and the thermal interface material (TIM) are included (Thermal via regions are shown in Figure 2.9).

In Figure 2.9, a schematic of top view of top silicon die and the thermal via regions are shown. We assumed the density of thermal via in each region is 50% which leads to 5% of the entire die. The area is meshed by 200μm element size (50×50 grids).

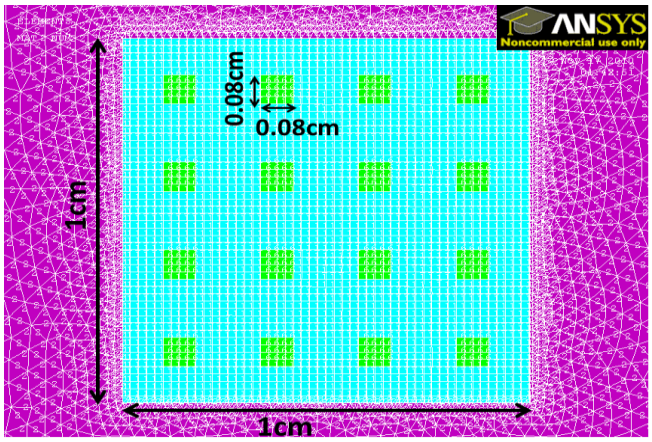


Figure 2.9. Top view of the meshed package IC structure. Thermal via regions are shown with green color in the figure ($0.08\times0.08cm^2$). Cyan is the Si die, and purple is the heat spreader.

Table 2.3. Material properties and dimensions of packaged model.

	Area (mm^2)	Thickness (mm)	Thermal Conductivity (W/m-K)	Density (kg/m^3)	Specific Heat (J/kg-K)
Si Die 1	10×10	0.2	148	2330	700
BCB	10×10	0.02	0.2	1051	2187
Si Die 2	10×10	0.5	117.5	2330	700
TIM1	10×10	0.025	5.91	1930	15
Heat Spreader	28×28	1	395	8933	397
TIM2	28×28	0.1	3.5	1100	1050
Heat Sink	60×80	4.5	395	8933	397

The power dissipation profiles in the top and bottom active layers are shown in Figure 2.10a and Figure 2.10b, respectively.

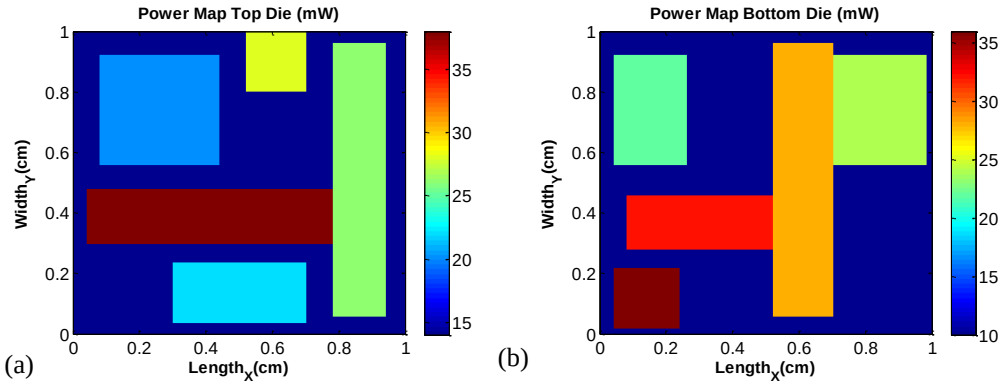


Figure 2.10. Power dissipation maps at the, (a) top active layer, (b) bottom active layer.

The temperature calculations using ANSYS and PB are shown in Figure 2.11. For the case of PB, it is important to include corrections using the method of images and also the intrinsic error compensation due to the heat spreading in larger heat sink in the package. In Figure 2.12, a comparison between these temperature profiles along the diagonal of the chip and also along the second row of the vias are presented. Filtered and non-filtered results are both shown and compared with ANSYS results. The significance of including intrinsic error compensation step, is also indicated in this figure. For this case study, including intrinsic error compensation in each layer lead to up to 10% less error in final temperature profiles. The relative temperature error and also computation times are presented in Table 2.4. PB results are obtained 76 times faster than ANSYS. The maximum error of 12% in the entire temperature profile relative to ANSYS is due to temperature difference of less than 2.3 degree (81.2-78.9), which happens at the boundary between via region and the Si chip.

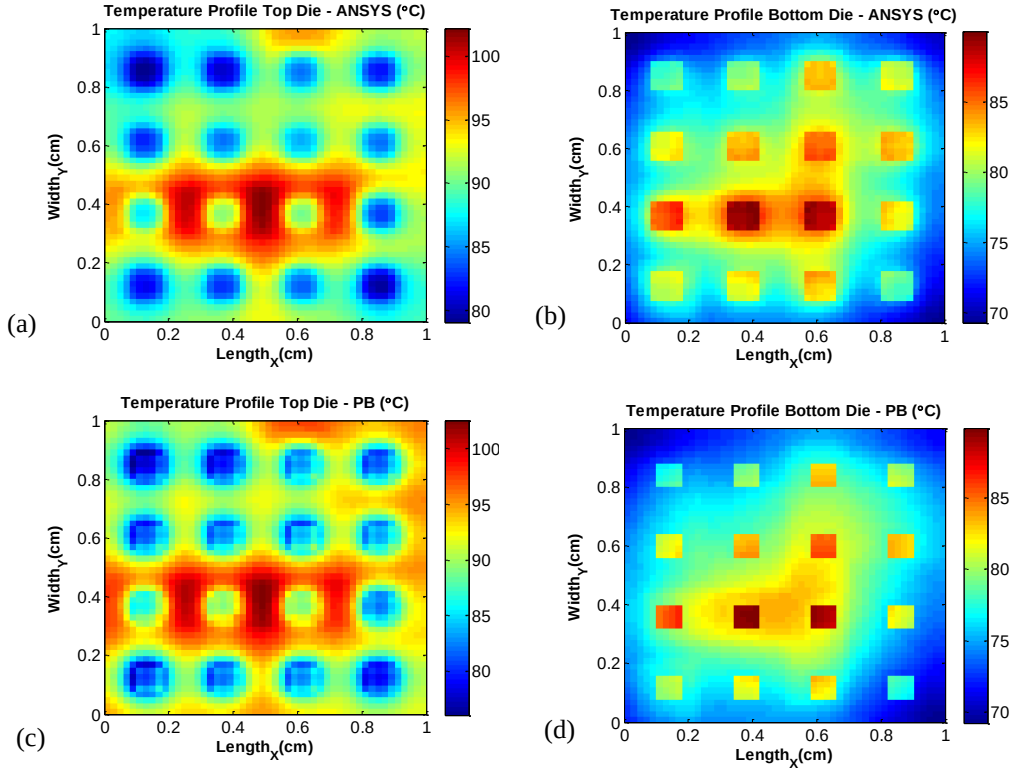


Figure 2.11. Temperature map, ANSYS (a) top, and (b) bottom active layer. Temperature map, PB (c) top, and (d) bottom active layer.

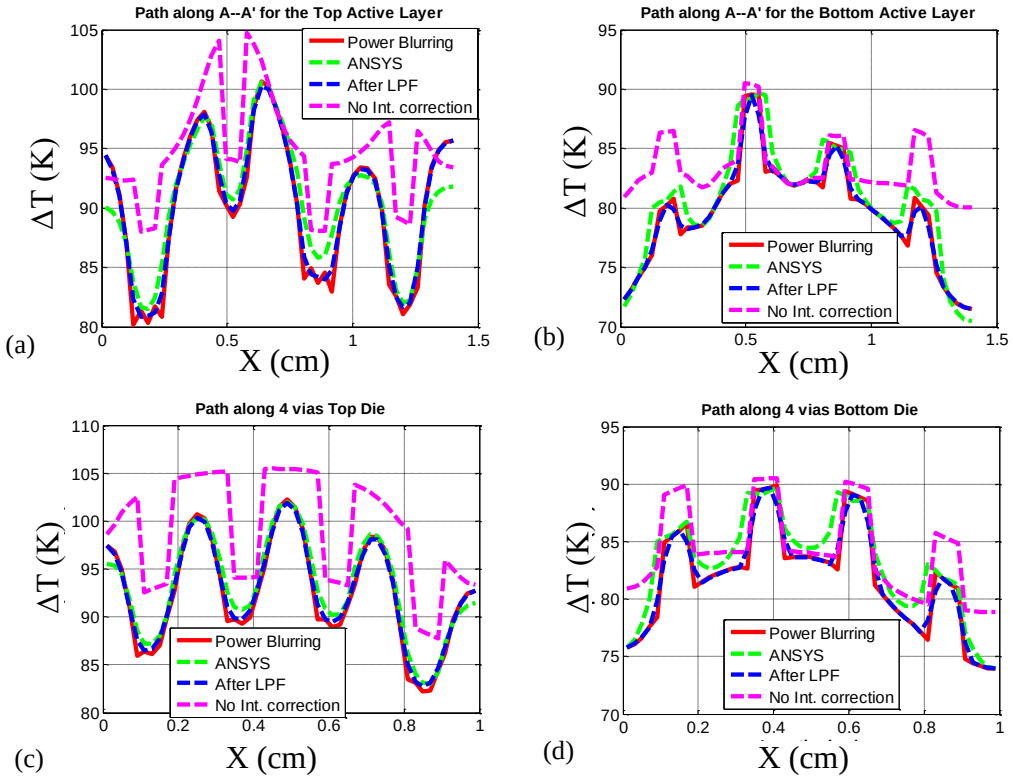


Figure 2.12. Comparison between temperature profiles along chip diagonal at the (a) top, and (b) bottom active layers. Comparison along 2nd row vias (c) top, and (d) bottom active layers.

Table 2.4. Detailed Comparison between PB and ANSYS.

	ANSYS	PB (Top)	PB (Bottom)
Max Err.	-	7.98%	12.14%
Err. In hottest spot	-	0.023%	0.58%
Execution time	91.1s	1.2s	-

Influence of implementing thermal via into 3D ICs is presented in Figure 2.13. Inclusion of thermal via regions resulted in ~%30 reduction in the maximum temperature at the top surface (Figure 2.13a and 2.13b). However, at the bottom surface and at the bottom of

thermal vias, temperature will increase (Figure 2.13c and 2.13d). This means thermal vias facilitate heat transport from the top die to the package, but because of limitation in spreading at the bottom junction, localized temperature will increase. Thermoelectric modules for localized solid-state cooling at the bottom of thermal vias is a possible solution to resolve this issue.

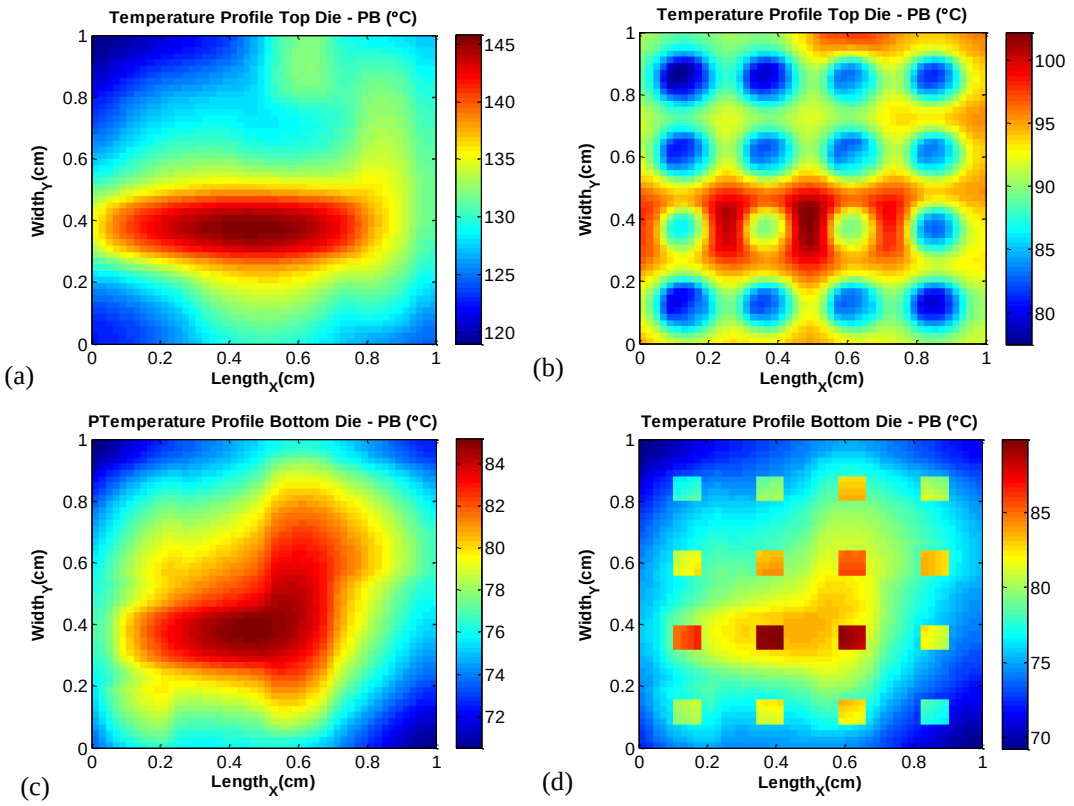


Figure 2.13. Comparison between temperature profiles between PB results in 3D ICs with and without implementation of thermal vias. Temperature profile at the top: (a) with via, and (b) without vias. Temperature profile at the bottom (c) with via, and (d) without vias.

2.5. Summary

The demands for efficient thermal simulation for VLSI ICs in a thermal package as well as power electronic and optoelectronic devices have risen as the CMOS technology scales down and power densities escalates. In this work, we presented a highly accurate yet fast thermal simulation method, namely Power Blurring (PB). We demonstrated that the PB method is suitable for both static and transient thermal simulations, while, unlike the conventional Green's function based methods, realistic package models can be incorporated. PB is able to solve nonlinear problems (calculating temperature profiles in ICs considering temperature dependence of material properties of the chip) using only two or three iterations. Moreover, the power blurring method is used to estimate the temperature profile in 3D ICs which include thermal vias. The method is validated against FEM (ANSYS), and the results are obtained 76x faster, while the error at the hot spot was less than %0.6. Power Blurring is able to incorporate any non-uniformity in the chip which can be due to thermal vias, or even chip with different material properties in different regions. This advantage becomes more prominent when one considers the fact that every FEA method need to re-mesh and recalculate the temperature profiles for new power distribution profiles while in the case of PB once the thermal masks for a given geometry are obtained, then for any power map the temperature profile can quickly be obtained. In particular, this is very beneficial at the early stage of IC design for placement and routing of the thermal via regions and functional blocks. Obtaining temperature profiles by performing convolution at each meshed element, based on material properties in that element, is another new idea that is used. This will increase the accuracy of resulted

temperature profile and versatility of power blurring at the expense of losing a bit of speed. The latter can be easily compensated by using a lower level programming (e.g. C) and with the use of dedicated graphical processing units (GPUs).

3. DESIGNING A MECHANICALLY ROBUST THERMOELECTRIC MODULE FOR HIGH TEMPERATURE APPLICATIONS

3.1.Introduction

For every unit of energy that is converted into electricity, more than one or two units of energy are not used in power plants today. This excess energy is primarily wasted as heat – or thermal energy. Deploying systems to recover this wasted heat back into the energy stream is a wide spread topic of research. Thermoelectric generators (TEGs) are emerging as a possible solution for high temperature energy conversion applications and waste heat recovery systems. The key challenges are improving the efficiency of thermoelectric power generator module (TEM) and its material cost in large scale production.

A system optimization for TE waste heat recovery system and minimization of the TEM cost is presented in [20]. The closed form analytical solution reveals that the optimum solution for the maximum output power can be obtained by both electrical and thermal impedance matching and together with their heat source and the heat sink (hot and cold reservoirs). Upon finding the optimum solution, cost-performance analysis is conducted to find the minimum cost design at a given system efficiency. This optimization elucidates that the fractional area coverage of the thermoelectric (TE) leg, called fill factor or *FF*, plays a significant role in minimizing the mass of the TE material used in thermoelectric waste heat recovery systems.

It is shown in [20] that improving the figure-of-merit (*ZT*) along with decreasing the fill factor would further reduce the total cost. Because the maximum power output from a thermoelectric system is proportional to square of temperature difference between the hot and cold reservoirs [75], employing a thermoelectric generator with optimum design in high temperature applications and with large temperature difference, such as on top of a steam turbine cycle, will be an economical approach to increase energy production [21]. However, both reduced fill factor and higher temperature range imply a larger impact on thermo-mechanical reliability. Elevated thermal stresses are viewed today as major bottle-necks for reliability and robustness in high temperature TEM applications. These stresses are caused, first of all, by the significant differences in temperature between the "hot" and the "cold" ceramic plates in a TEM design (Figure 3.1). The thermal stress problem can be solved by selecting adequate thermoelectric materials [76], [77] as well as by finding effective ways to reduce the stress level [78].

In this Chapter, an analytical and a finite-element-analysis (FEA) models are used to evaluate the thermal stresses in a simplified (two-leg) TEM design. We will demonstrate that, in the case of simplified TEM, by reducing fill factor as well as using compliant interface materials, one can reduce the maximum shearing stress occurring at the contacts. The maximum shear stresses are supposedly responsible for the structural robustness of the TEM assembly [78]. It should be pointed out, that reduction of the maximum shear stress by decreasing fill factor is not universal and it depends on other parameters, such as the coefficients of thermal expansion (CTEs) of the different layers and the structural boundary conditions for the TEM assembly. Therefore, we exploit finite element analysis to study

how the geometry, structural boundary conditions, and the CTE of the materials in the TEM structure would change the maximum shear stress particularly in a high temperature application.

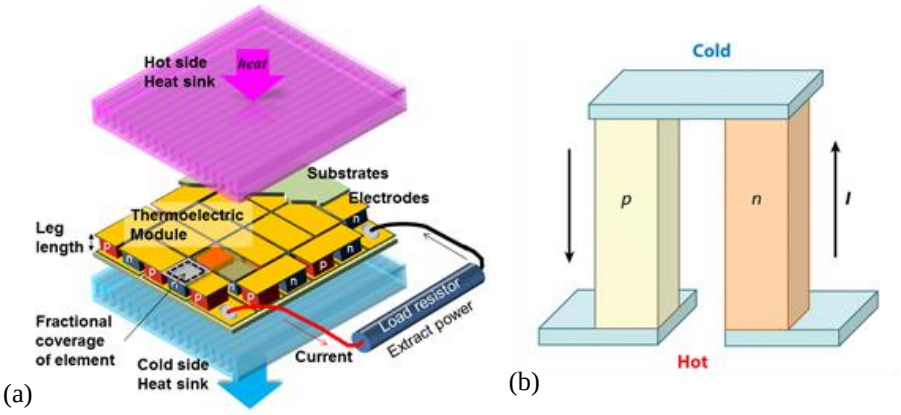


Figure 3.1. Thermo-electric Module; (a) General view; (b) A two leg module with n-type and p-type legs.

The rest of this chapter is organized as follows. The analytical model is described in section 3.2. In section 3.3, we have employed the model to calculate the shearing stress in different TEM designs. Discussion of the results and comparison with FEM data are also presented. Moreover, we investigate the impact of variation in the geometry, boundary conditions, and the coefficient of thermal expansion (CTE) of the materials on the maximum shearing stress in a thermoelectric power generator module (TEM) proposed for high temperature applications. The chapter concludes in section 3.4 with a summary and possible future work.

3.2.Analytical Modeling

3.2.1. Assumptions

The following major assumptions are used in the analysis:

- All the materials behave in the elastic fashion.
- Instead of addressing the actual three-dimensional TEM structure, a two-dimensional longitudinal cross-section of this structure idealized as a long-and-narrow strip could be considered.
- The bonded TEM ceramic components can be treated, from the standpoint of structural analysis, as elongated rectangular plates that experience linear elastic deformations, and approximate methods of structural analysis and materials physics can be used to evaluate the induced stresses and displacements;
- The interfacial shearing stresses can be evaluated based on the concept of the interfacial compliance [79].
- The interfacial compliances of the bonded components and the TEM legs can be evaluated, however, based on the Rebière solution in the theory-of-elasticity for a long-and-narrow strip (see, e.g., [79]).
- The assembly is thick and stiff enough, so that it does not experience bending deformations, or, if it does, bending does not affect the interfacial thermal shearing stresses and need not be accounted for.
- The interfacial shearing stresses can be evaluated without considering the effect of "peeling", i.e., the normal interfacial stresses acting in the through-thickness direction of the assembly.

- The longitudinal interfacial displacements of the TEM bonded components can be sought as the sum of 1) the unrestricted stress-free displacements, 2) displacements caused by the thermally induced forces acting in the cross-sections of the TEM components and 3) additional displacements that consider that, because the thermal loading is applied to the component interface, the interfacial displacements are somewhat larger than the displacements of the inner points of the component.
- TEM legs provide mechanical supports for the TEM bonded components (ceramics) and their interfacial compliance is critical when one intents to buffer the interfacial stress, but do not experience thermal loading themselves.

Some additional, more or less minor, assumptions are indicated in the text of the chapter.

3.2.2.Interfacial Compliance

Analytical modeling uses the interfacial compliance concept suggested in Refs. [79]–[81]. The concept enables one to separate the roles of the design (its geometry and material properties) and the loading caused by the change in temperature and/or temperature gradients. The approach is based on and reduced to the evaluation of the longitudinal interfacial compliance of a strip subjected to the longitudinal shear loading applied to its long edge (Figure 3.2a). An important assumption underlying the rationale behind the employed analytical model is that the actual 3D structural element (experiencing in a multi-material body interfacial loading caused by the dissimilar materials in the body) can be substituted by an elongated strip that is, in effect, the longitudinal cross-section of the body. The following approximate formula for the longitudinal displacements of the edge of such

a strip has been used ([79]–[81]) to evaluate the longitudinal displacements of a strip loaded over its long edge by a distributed shear loading:

$$u_0 = -\frac{1-\nu^2}{Ehb}\int_0^x Q(\xi)d\xi + \kappa\tau_0(x) \tag{3.1}$$

Here E and ν are the modulus of elasticity and Poisson's ratio for the strip material, κ is the longitudinal compliance of the strip (defined as the ratio of the longitudinal displacement to the loading $\tau_0(x)$, h is the thickness of the strip, b is its width, and $Q(x)$ is the distributed longitudinal force acting at the x cross section of the strip. The first term in the equation (1) reflects an assumption that the displacement of the strip's edge at the x cross section is uniformly distributed over the cross section. The second term account for the deviation of the actual, non-uniform, distribution of this force: the longitudinal displacements at the strip edge, where the load $\tau_0(x)$ is applied, are somewhat greater than at the inner points of the cross section. The structure of this term reflects an assumption that the correction in question can be calculated as the product of the shearing load $\tau_0(x)$ in the given cross section and the longitudinal compliance of the strip, as well as an assumption that the displacement determined by this term is not affected by the states of stress and strain in the adjacent cross sections. The detailed rationale behind the formula (1) and the subsequent derivation of the interfacial compliance κ can be found in Refs. [79]–[81]. The obtained general formula for this compliance is:

$$\kappa = \frac{\sum_k \gamma_k M(u_k) sin\alpha_k x}{Eb\sum_k \alpha_k \gamma_k sin\alpha_k x} \tag{3.2}$$

Here the function $M(u_k)$ and the parameters α_k , u_k and γ_k are defined as:

$$M(u_k) = \left(\frac{1+\nu}{2}\right)\left[(3-\upsilon-(1+\upsilon)u_k cotanhu_k)cotanhu_k + (1+\upsilon)u_k - \left(\frac{2(1-\nu)}{u_k}\right)\right]$$

$$\alpha_k = \frac{k\pi}{2l}, \quad u_k = \alpha_k h = \frac{k\pi}{2}\frac{h}{l}, \quad \gamma_k = \frac{2}{\alpha_k l}\int_0^l \tau_0(x) sin\alpha_k x\, dx, \quad k = 1,3,5,7,\ldots \qquad (3.3)$$

Only the odd numbers are used in the formulas (3.2) and (3.3), because the strip deformations are symmetric with respect to its mid-cross-section.

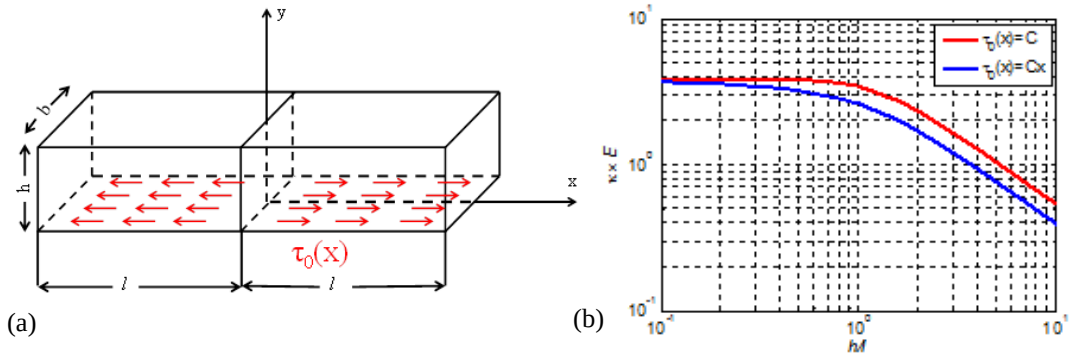


Figure 3.2. Interfacial compliance in a stipe. (a) Elongated strip subjected to shear loading. (b) Evaluation of longitudinal interfacial compliance coefficient for linear shear load distribution (blue curve) as well as uniform shear load distribution (red curve).

The interfacial compliance κ depends on the geometry of the strip (ratio *h/l* of its height *h* to half of its length *2l*), elastic constants of its material, and, as shown in Figure 3.2b, also, slightly, on the shear load $\tau_0(x)$. The latter effect is, however, insignificant in comparison with the effects of the aspect ratio of the strip and the properties of its material and in an approximate analysis could be neglected. The expression (3.2) can be approximated, in extreme situations, by simplified relationships and, in the cases of extreme aspect ratios *h/l*, leads to the following simple formulas:

$$\kappa = \begin{cases} \frac{h}{3Gb} & \frac{h}{l} < 0.5 \\ \frac{3-\nu}{2\pi b}\frac{l}{G} & \frac{h}{l} > 2 \end{cases} \tag{3.4}$$

Here $G = \frac{E}{2(1+\nu)}$ is the shear modulus of the material. The compliance for the intermediate *h/l* ratios can be obtained, in an approximate analysis, by interpolation. In our further analysis we use, however, the formulas (3.4) when the *h/l* ratio is below 0.5 or above 2, and the general formula (3.2), when the *h/l* ratio is between 0.5 and 2.0. The κ values computed for different materials employed in the TEM under evaluation (Figure 3.3), are based on the assumption that the loading $\tau_0(x)$ is uniformly distributed over the long edge of the strip, i.e., coordinate *x* independent.

3.2.3.Shearing Stress in Two-leg TE Module

As has been indicated above, the analysis is conducted under the major assumption that the bonding systems ("legs") provide mechanical support in the TEM design and their interfacial compliance, in terms of providing a strain buffer between the TEM components is important, but do not experience thermal loading themselves. This assumption seems to be justified in the case of short enough bonds, i.e. in the case of long assemblies with short bonded regions which is the primary situation of interest in this analysis. Such an assumption might result, however, in an overestimation of the induced stresses in the case of not-very-short bonded regions ("not very thin legs"), but could still be supposedly used for the relative assessment of the state of stress in a TEM design in question.

The longitudinal interfacial displacements can be predicted using the following approximate formulas:

$$u_1(x) = -\alpha\Delta t_1 x + \lambda_1 \int_0^x T(\zeta)d\zeta - k_1\tau(x); u_2(x) = -\alpha\Delta t_2 x - \lambda_1 \int_0^x T(\zeta)d\zeta + k_1\tau(x) \quad (3.5)$$

These formulas are similar to those used in [81], where, however, dissimilar bonded component materials were considered. The first terms in the equations (3.5) are unrestricted thermal expansions of the TEM components. In these terms, α is the coefficient of thermal expansion (CTE) of the material, Δt_1 and Δt_2 are the change in temperature of the components (from the manufacturing temperature to the operation temperature). The second terms are due to the axial thermally induced forces *T(x) = Q(x)/b* acting in the cross section of the components. In these terms, $\lambda_1 = \frac{1-\upsilon_1}{E_1 h_1}$ is the axial compliance of one of the bonded components, E_1, υ_1, h_1 are the elastic (Young's) modulus, Poisson's ratio of the material and the component thickness, respectively. These terms were evaluated using Hooke's law. The last terms in equation (3.5) account for the actual non-uniform distribution of the forces *T(x).* These terms reflect an assumption that the corresponding corrections can be evaluated as products of the interfacial compliance and the interfacial shearing stress acting in this cross-section. It is assumed also that these terms are not affected by the states of stress and strain in the adjacent cross-sections.

The condition of the compatibility of the interfacial displacements of the bonded TEM components can be written as

$$u_1(x) = u_2(x) + \kappa_0\tau(x) \quad (3.6)$$

where κ_0 is the interfacial compliance of the buffering material (structure). Substituting the displacements (3.5) into the condition (3.6), one can obtain a governing equation for the force *T(x),* and the solution to this equation can be sought in the form:

$$T(x) = \frac{-\alpha \Delta T}{2\lambda_1} + C_1 sinh\, kx + C_2 cosh\, kx \tag{3.7}$$

where the first term is the particular solution for the inhomogeneous governing equation and the second and the third terms provide the general solution to the corresponding homogeneous equation. The constants C_1 and C_2 are constants of integration.

$$k = \sqrt{\frac{2\lambda_1}{\kappa}} \tag{3.8}$$

In equations (3.7) and (3.8), k is the parameter of the interfacial shearing stress and κ is the total interfacial compliance of the TEM assembly.

The solution (3.7) must satisfy the boundary conditions:

$$T(-l) = \hat{T}, T(l) = 0 \tag{3.9}$$

The compatibility condition for the longitudinal displacements at the bonded and unbonded regions of the TEM assembly can be written as follows:

$$\kappa\tau(-l) = 2\lambda_1 \hat{T}(L - 2l) \tag{3.10}$$

In these equations, $\hat{T}$ are the forces that determine the role of the global mismatch of the components. Global mismatch occurs outside the bonded region because of the mismatch of assembly components, while local mismatch occurs within the bonded regions and is due to the thermal mismatch of the materials. From equations (3.7), (3.8), (3.9) and the obvious relationship $T'(x) = \tau(x)$, one can find the force $\hat{T}, and\, T(x)$ acting in the unbonded and bonded regions, respectively, and the interfacial shearing stress. Subsequently, the following expression for the interfacial shearing stress could be obtained [78]:

$$\tau(x) = k\frac{\alpha\,\Delta T}{2\lambda_1}\left[\frac{sinh\,kx}{cosh\,kl} + \frac{tanh\,kl}{2kl\left(\frac{L}{2l}-1\right)sanh\,2kl+cosh\,2kl}cosh\,k(l-x)\right],\ \frac{L}{2l} \geq 1 \qquad (3.11)$$

The maximum interfacial shearing stress takes place at the assembly edges:

$$\tau(l) = k\frac{\alpha\,\Delta T}{2\lambda_1}tanh\,kl\left[1 + \frac{1}{2kl\left(\frac{L}{2l}-1\right)sanh\,2kl+cosh\,2kl}\right],\ \frac{L}{2l} \geq 1 \qquad (3.12)$$

This relationship indicates that by decreasing the product *kl* of the parameter of the interfacial shearing stress and half the length of the bonded region one could reduce the maximum interfacial shearing in this region.

3.2.4.Shearing Stress in Multileg TE Module

We used the approach proposed in [82], to develop analytical equations (Equation 3.13) for a simplified six leg model, shown in Figure 3.12b, and to calculate maximum shear stress in the structure.

$$\tau_{max} = k\frac{\alpha\Delta t}{\lambda_1}\tanh kl\left[1 + \frac{(\tanh kl + \coth kl + 8kl\left(\frac{L}{2l}-1\right))}{(\left(\tanh kl + \coth kl + 4kl\left(\frac{L}{2l}-1\right)\right)^2 + 8kl\left(\frac{L}{2l}-1\right)\tanh kl + (4kl\left(\frac{L}{2l}-1\right))^2)\ \sinh 2kl}\right] \qquad (3.13)$$

In these equations, *k* is the parameter of the interfacial shearing stress (equation 3.8), κ is the total interfacial shear compliance of the mid-layers between the two top and bottom components, *α (1/˚C) ,v_1, E_1 (GPa),* and h_1 *(m)* are the coefficient of thermal expansion (CTE), Poisson ratio, modulus of elasticity, and the thickness of the substrate component, respectively, *Δt* is temperature difference between the hot and the cold sides, $\lambda_1 = \frac{1-v_1}{E_1 h_1}$ is the axial compliance of one of the bonded components, *L* is end to end distance between

the two legs, and *l* is the half width of thermoelectric leg. Major assumptions in obtaining the equations are listed in section 3.2.1.

3.3.Case Studies

3.3.1.Two-leg simplified TE Module

The analytical solution described in section 3.2 is applied to the TEM structure shown in Figure 3.3. The material properties are given in Table 3.1. Three different assembly sizes of 10mm, 20mm and 40mm (L=5mm, 10mm, and 20mm) were chosen. The value of *l*, the half the bonded region length, has been varied to evaluate its effect on the maximum interfacial shearing stress. The temperature difference between the top and the bottom components (ΔT) is 130ºC. The thickness of the TEM leg is 4mm. Bonded region and component thicknesses are as indicated in Figure 3.3.

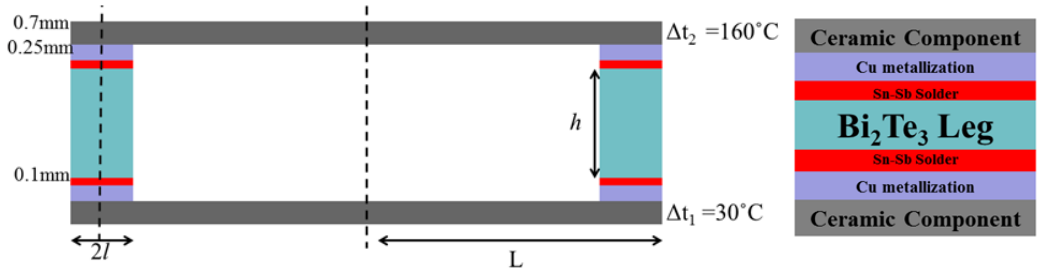


Figure 3.3. Thermo-Electric Module 2D structure. Parameters shown in the figure are used for case studies in section 3.3.

The maximum interfacial shearing stress versus bonded region length for different assembly lengths is obtained using the analytical model presented in section 3.2. The results are plotted in Figure 3.4. As it can be seen, the decrease in the length of the bonded region results in lower maximum interfacial shearing stress. Also, for the same bonded

region length, the increase in the assembly's length leads to the decrease in the maximum interfacial shearing stress. This means that the increase in the L/*2l* ratio and the decrease in the fractional area coverage of the thermoelectric legs lead to lower maximum interfacial shearing stresses.

Finite Element Modeling (FEM) software, ANSYS, has been used to simulate the same TEM assembly. 8 nodes plane 223 elements in plane strain mode were used. The structure is meshed with very fine square elements. Each element is 25x25μm^2 and there were around 400000 elements in this structure. The boundary conditions of the simulations were set according to the boundary conditions in the analytical model. The strain free temperature (reference temperature) is set to zero. Then the top component heated up to 160 ºC and the bottom component is heated up to 30ºC. Temperature load applied to the structure is shown in Figure 3.5a. The deformed shape of the TEM structure due to this temperature load is superimposed on the edge of original un-deformed model. The translation is restrained in both x and y direction in the bottom corners of the TEM structure. A vertical cross section of the temperature profile is plotted in Figure 3.5b. The material properties are set according to Table 3.1. The coefficient of thermal expansion (CTE) for the ceramic plates is set to 6.5×10^{-6} (1/ºC).

Table 3.1**.** Mechanical properties of materials employed in TEM

Material	**Young Modulus (GPa)**	**CTE (ppm/ ºC)**	**Poisson's ratio**
Ceramic Component	380	6.5	0.28
Copper Stripe (metallization)	115	17	0.31
Sn-Sb solder layer	44.5	27	0.33
Be_2Te_3 leg	47	16.8	0.4

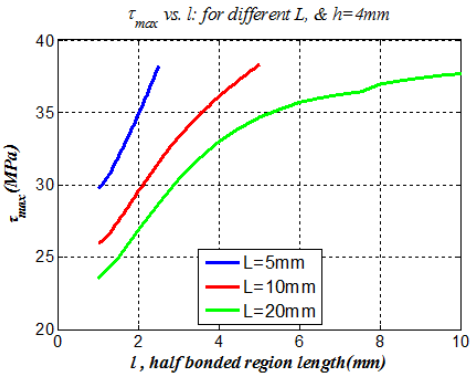
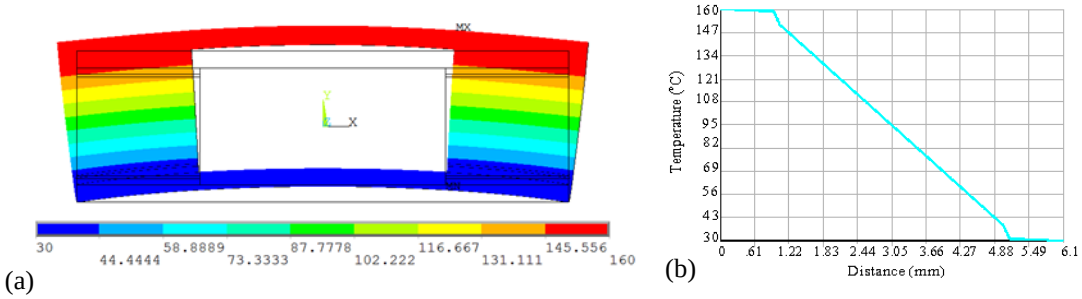


Figure 3.4. Variation of maximum interfacial shear stress (τmax) vs bonded region's length. Different assembly lengths considered: 2L= 10mm (blue), 20mm (red), and 40mm (green).

For other layers we consider two cases. In the first case we set the same CTE for them as the components. In the second case we set them to the value shown in Table 3.1.



(a) (b)

Figure 3.5. Thermoelectric Module deformed shape due to high temperature. (a) A sample deformed shape of a 2D TEM simulated in ANSYS is superimposed on the edge of the un-deformed 2D TEM Model. (b) Temperature along vertical cross section of one of the legs.

The maximum shearing stresses for different assembly lengths and bonded region length sizes are calculated using ANSYS. The results are compared with the analytical solutions in Figure 3.6a, and 3.6b. In Figure 3.6a the half assembly length (L) is 5mm, the leg thickness (*h*) is 4mm and half bonded region length (*l*) is changing 0.5mm to 2mm. Also, the same simulations have been done for an assembly with L equal to 10mm and *l* changing

from 1mm to 4mm. It can be seen in both cases that our analytical model follows the same trend as FEM and the results are in good agreement. Changing fractional coverage area by a factor of 16 from 64% to 4% resulted in a maximum 40% drop in the interfacial shear stress in each case. Figure 3.7 shows that if we consider CTE mismatch between the layers as well the values for stress would increase significantly for this particular case study. However, this Figure is also indicative of the same trend for maximum interfacial shear stress. It can be seen in Figure 3.7 that, for the TEM structure used in this case study with the material property listed in Table 3.1, lowering the fractional coverage area by a factor of 16 from 64% to 4% will lead to about 32% drop in maximum interfacial shear stress, which matches our analytical model.

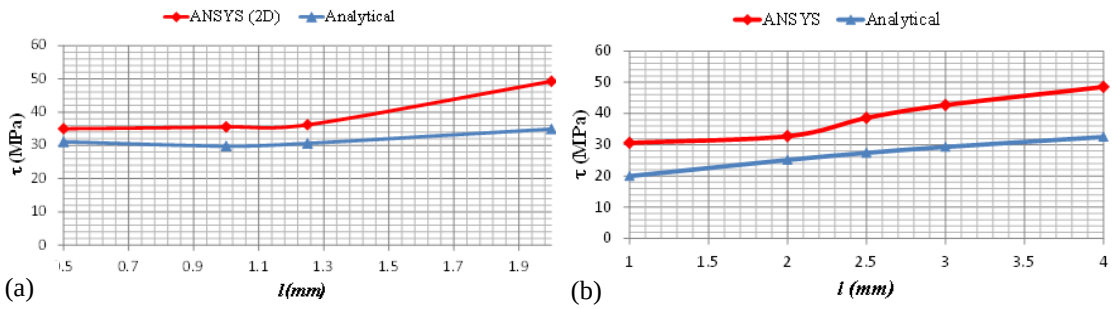


Figure 3.6. Maximum interfacial shear stress against TE leg length (ANSYS vs Analytical). (a). Maximum interfacial shear stress obtained by ANSYS (Red curve) and Analytical model (Blue curve) for h=4mm and L=5mm. (b). Maximum interfacial shear stress obtained by ANSYS (Red curve) and Analytical model (Blue curve) for h=4mm and L=10mm.

The decrease in the TEM leg thickness results in higher maximum interfacial shearing stresses. This is shown in Figure 3.8. Both ANSYS and the analytical model show similar results as the thermoelectric leg thickness decreases. These simulations are performed for

a structure with L and *l* equal to 5mm and 2mm, respectively. Increasing the leg thickness by a factor of 3.8 leads to a 70% drop in the maximum interfacial shearing stress. As evident from these figures, analytical and numerical data are in good agreement.

It is indicated in Figure 3.4 that by decreasing the bonded region length the maximum shear stress decreases. On the other hand, by decreasing the TE leg thickness, the maximum shear stress would increase. Therefore, employment of thinner and longer legs could indeed result in a substantial stress relief, thereby leading to a more mechanically robust TEM. In [76] a similar conclusion was achieved with 3D simulation of a 2 leg thermoelectric module.

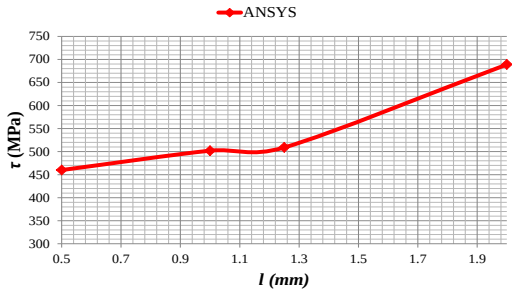


Figure 3.7. Maximum interfacial shear stress. ANSYS simulation result (Red curve) for the TEM structure with right values of CTE (shown in Table 3.1).

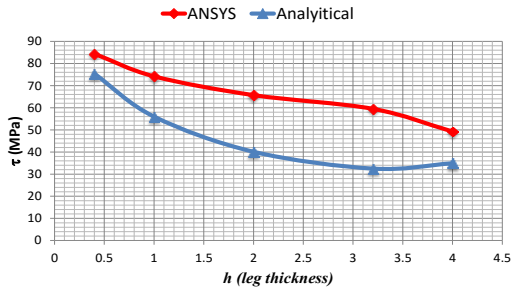


Figure 3.8. Maximum interfacial shear stress ANSYS vs Analytical for different thicknesses. ANSYS (Red curve) and Analytical model (Blue curve) for L=5mm and *l*=2mm, while h changing from 0.4mm to 4m.

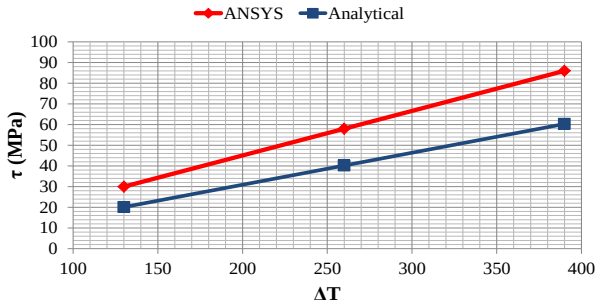


Figure 3.9. Maximum interfacial shear stress vs. top and bottom components temperature difference (ANSYS vs Analytical). ANSYS (Red curve) and Analytical model (Blue curve) for L=10mm and *l*=1mm, and h=4m.

Equation (3.12) shows that, based on our analytical model, the maximum interfacial shear stress varies linearly with the temperature difference between the top and bottom ceramic components. Three simulations are performed in ANSYS with ΔT equal to 130˚C, 260˚C, and 390˚C. For these simulations we chose a 2D structure with L, *l*, and h equal to 10mm, 1mm, and 4mm, respectively. The results are plotted in Figure 3.9. As it can be seen in the figure, increasing of the temperature difference between the components by a factor of 2 and 3 will result in a factor of 2 and 3 augmentations in the value of maximum shear stress in both the analytical model and ANSYS results.

3D simulation is also carried out to confirm what was obtained analytically. The meshed structure is shown in Figure 3.10. Symmetry is used and a quarter of the model is simulated. Again, the simulations are performed for two half assembly lengths (L) of 5mm and 10mm. The TE leg thickness is chosen to be 4mm. The temperature difference between the ceramic components is set to 130 ˚C. By changing the bonded region length, the fractional coverage area in both cases is reduced from 64% to 4%. As can be observed in Figure 3.11, 3D simulation shows that the maximum shear stress reduces by a maximum 80%. A

comparison between the results of 2D ANSYS simulation, analytical model, and 3D ANSYS simulation for both case studies is shown in Figure 3.11. The analytical solution has occurred between the two FEA solutions and follows the same trend. Intuitively this behavior seems reasonable, as the analytical model indirectly, by bringing in the Young's modulus and the Poisson's ratio, takes into account (in an approximate fashion) the 3D state of stress. The difference between the maximum shear stress values, obtained by analytical model in comparison with ANSYS 2D and 3D results, is due to the fact that the analytical model is neither a 3D nor an exactly 2D model. The reason is the geometry of TEMs is complex and cannot be considered as a plane strain or plane stress problem. Our Quasi-2D analytical model takes the Poisson's ratio into account (in an approximate fashion), while in ANSYS and other FEM a plain strain or plane stress condition for 2D problem needs to be defined.

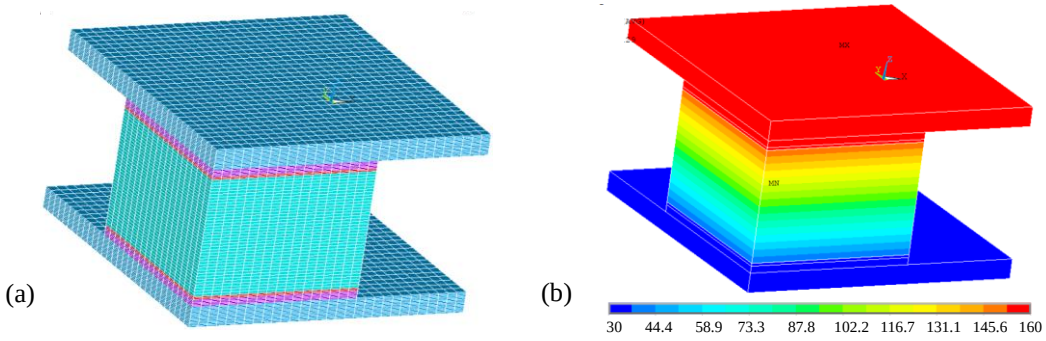


Figure 3.10. 3D Modeling ANSYS. (a) 3D meshed structure. (b) Temperature profile.

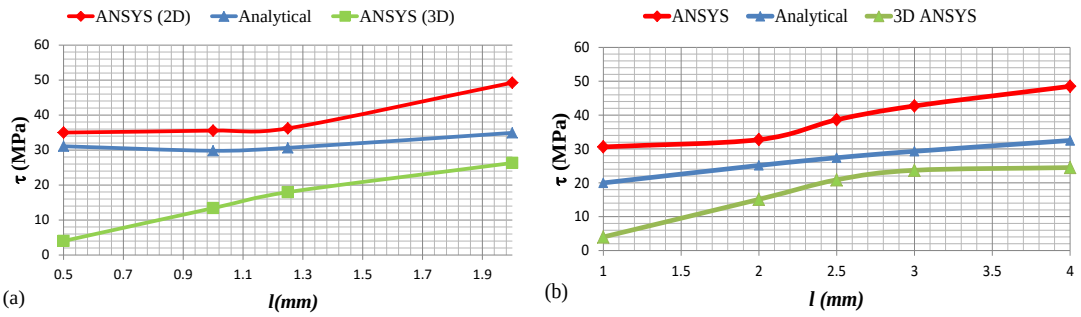


Figure 3.11. Comparison between Analytical results with 3D and 2D ANSYS results. Maximum shear stress vs. half bonded length is plotted for (a) L=5mm, h=4mm. (b) L=10mm, h=4mm.

3.3.2.Multileg High Temperature TE Module

A schematic of a thermoelectric module proposed for high temperature application is shown in Figure 3.12 [82]. The substrate components and metallization layers are made of molybdenum (Mo) alloys. The rest of material properties and dimensions are listed in Table 3.2. The pitch distance between legs is set to 200μm. By changing the width of the TE legs, the fill factor can be changed. Equation 3.13 can be used for the simplified six leg model, shown in Figure 3.12b, to calculate maximum shear stress in the structure.

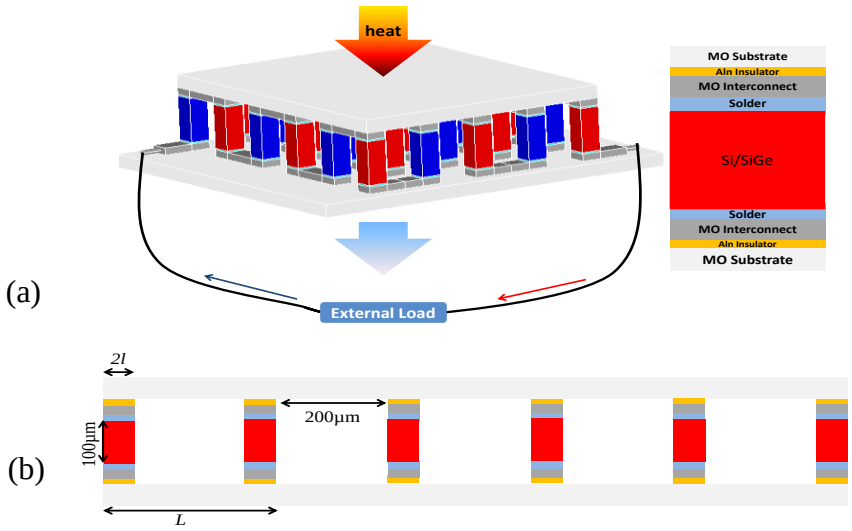


Figure 3.12. Multileg TE module structure. (a - left) Schematic of a Thermoelectric Generator, (a-right) and Materials proposed for different layers for high temperature applications [82]; (b) Simplified 2D model used for analytical modeling.

Table 3.2. Material properties and dimension for the proposed TEM for high temperature

Property / Material	Thickness (μm)	Young Modulus (Gpa)	CTE (10^{-6}/°C)	Poisson Ratio	Yield Stress (GPa)	Ultimate stress (GPa)
Mo Alloy	50	330	4.8	0.31	-	-
AlN Insulator	0.05	330	6.58	0.24	-	-
Mo Interconnect	30	330	4.8	0.31	-	-
Solder	5	78.5	14.2	0.42	200	220
Si/SiGe TE	100	250	2.6	0.28	-	-

The maximum shear stress takes place at the end of the peripheral legs. Maximum shear stress as a function of fill factor is graphed in Figure 3.13, assuming temperature at the hot and the cold sides are 800^0C and 630^0C (Δt=170). A 2x reduction the in maximum shear

stress is obtained by decreasing fill factor from 25% to about 3%. The maximum shear stress is plotted for two different values of CTE for substrate layer, so that we can compare the results with the corresponding numerical results in the following.

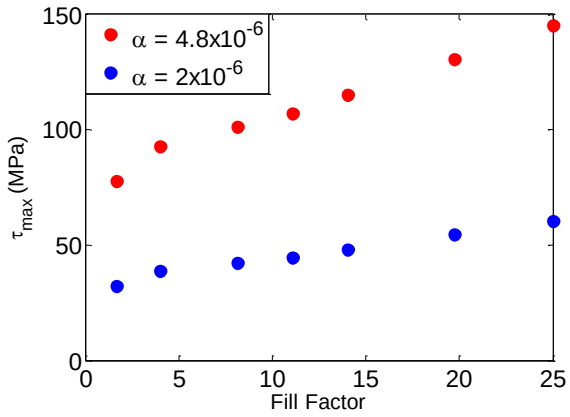


Figure 3.13. Analytical modeling result for maximum shear stress as a function of fill factor in variations of α. α is the CTE of Mo Substrate.

Although analytical modeling gives an intuition on the parameters contributing to the maximum shear stress, it has some key assumptions which impose limitations on the accuracy of the calculated thermal stresses. It assumes a homogenous CTE in all the layers and does not consider any local CTE mismatch between the layers. The analytical model takes into account the three dimensional (3D) state of stress in an approximate fashion by bringing in the Poisson ratio and elastic modulus of the substrate layer. Additionally, it assumes that the assembly is thick and stiff enough, so that it does not experience bending deformations, or, if it does, bending does not affect the interfacial thermal shearing stresses and does not need to be accounted for. We exploited ANSYS to carry out 3D finite element analysis and identify how each of these unaccounted parameters would contribute to the

maximum interfacial shear stress in the assembly. The additional case studies in this chapter are intended to illustrate the impact of variation in the geometry, boundary conditions, and the coefficient of thermal expansion (CTE) of the materials on the maximum shearing stress in a thermoelectric power generator module (TEM) for high temperature applications.

Three-dimensional (3D) structural-thermal analysis is carried out in ANSYS to calculate thermal stresses in TEMs. Twenty node Solid226 tetrahedral elements are used for meshing of the structure. The material properties are according to the Table 3.2. Homogenous CTE equal to 2×10^{-6} is assumed among the layers, unless otherwise stated. We performed all the simulations with three different Fill Factor of 25%, 11% and 4%, corresponding to leg widths of 100μm, 50μm, 25μm, respectively. The pitch distance between the legs is set to be 200μm. Also, temperature at the hot and the cold sides are assumed 800^0C and 630^0C (Δt = 170K).

A two leg simple 3D model for the TE module is constructed and numerical analysis is conducted. The maximum shear stress takes place at the interface with the Mo substrate. This maximum shear stress, at the top and bottom Mo/AlN interfaces, is plotted against Fill Factor for three different boundary conditions in Figure 3.14a and 3.14b. Free-standing structure, and constraint on the perpendicular translation at either the hot, or the cold side, are the three boundary conditions considered. 1.5x to 2.5x reduction in the maximum shear stress by decreasing the fill factor by 8x is obtained for different boundary conditions, which confirms the results of analytical modeling shown in Figure 3.13. Stress values for the 3D model are lower than those of obtained with 2D analytical model which is also

reasonable. The maximum shear at the Mo/AlN interface is due to the rigidity and the large modulus of the insulator material. In practice this large stress can be avoided by utilizing a thin compliant interface between the two layers. Therefore, the main concern is the maximum shear stress at the interconnect/solder interface. In Figure 3.14c and Figure 3.14d, the maximum shear stress at the top and bottom interconnect/solder interfaces, for the same structure under the same boundary conditions, are plotted against Fill Factor. It is evident from these figures that the maximum shear stress is reduced as Fill Factor decreases. In all the cases, the maximum shear under the free-standing boundary condition is lower than the other cases.

When we impose constraint on perpendicular translation at the hot surface, the maximum shear at the top interface (closer to the hot side) is larger while it stays almost the same on the cold side. Limiting the expansion of the hot side, generates stronger stresses on this side. However, the stress can be relaxed along the legs towards the cold side, resulting in similar stress values compared with the unconstrained case. On the contrary, when we impose constraint on the cold interface, the maximum shear stress took place at the bottom interface (closer to the cold side) and the stress on the top interface remains similar to the free-standing case.

It is apparent from Figures 3.13a-d that the two cases of imposing constraint at the hot and the cold sides are complementary to each other, and studying one of them would be sufficient to understand the other. Therefore, in the rest of analysis we only show the result for the free-standing case as well as the case with perpendicular constraint on y-translation at the cold side. Also, since interconnect/solder interface is more prone to thermo-

mechanical failure, due to solder's low yield stress, we will show the results for maximum shear at this interface.

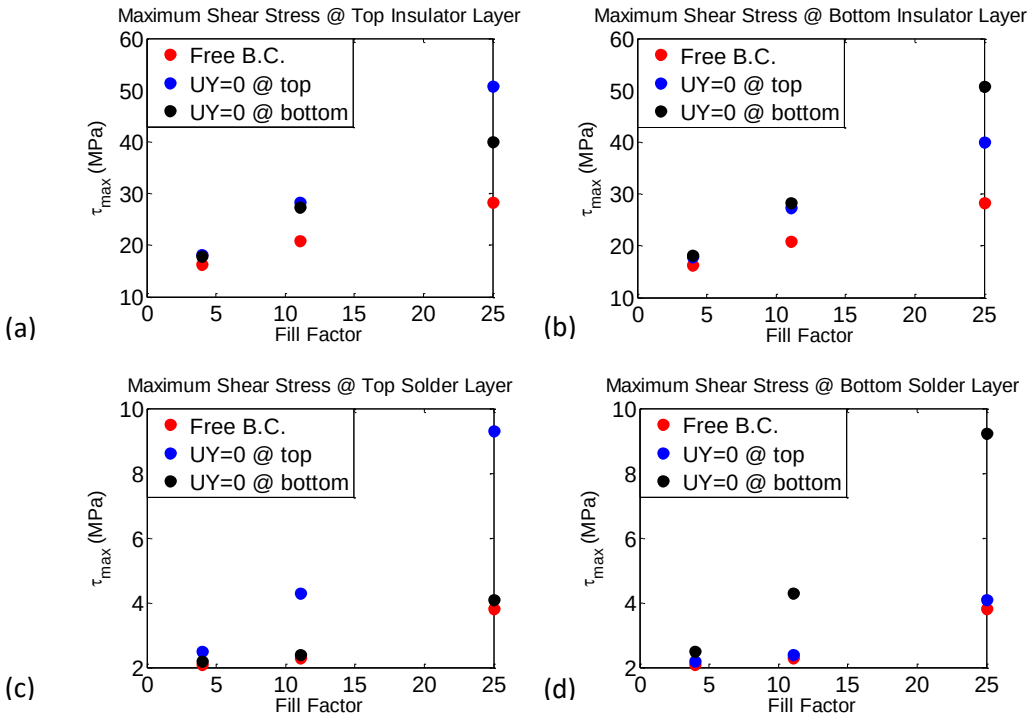


Figure 3.14. Maximum shear stress against fill factor for three different boundary conditions. Free-standing (blue), perpendicular translation is constrained at the top hot side (red), perpendicular translation is constrained at the bottom cold side (green). Maximum shear stresses are probed (a) at the top component/insulator interface; (b) at the bottom component/insulator interface; (c) at the top interconnect/solder interface; (d) at the bottom interconnect/solder interface.

In order to understand thermal stress behavior in a more realistic configuration, we performed three dimensional finite element analysis on 6-legs as well as an array of 6×6 legs (36-legs) structures (Figure 3.12a). Again, we are interested in the maximum shear stress at the interface of interconnect and the solder layer. Figure 3.15 reveals how the trends of maximum shear against fill factor vary for different cases of 2-legs, 6-legs, and 36-legs structures and different boundary conditions. For the case with free-standing

boundary condition, the maximum shear for a 2-legs simplified model would drop by reducing the fill factor, while the trends for the cases of 6 legs and 36-legs reverse and the maximum shear rises by decreasing the fill factor. For large fill factors, the maximum shear stress for the three structures is not different. However, at small fill factors the maximum shear stress is significantly different between 2-leg and multi-leg structures. This is mainly due to the asymmetric bending of the two large substrates. The outward/inward expansion of the hot and the cold substrates will deform the thermoelectric leg of the low fill factor structure and generate strong shear stresses.

If we anchor the bottom surface to limit the perpendicular translation at the cold side, the expansion of the bottom substrate would be limited and trends would be the same for different structures. This is shown in Figure 3.15b.

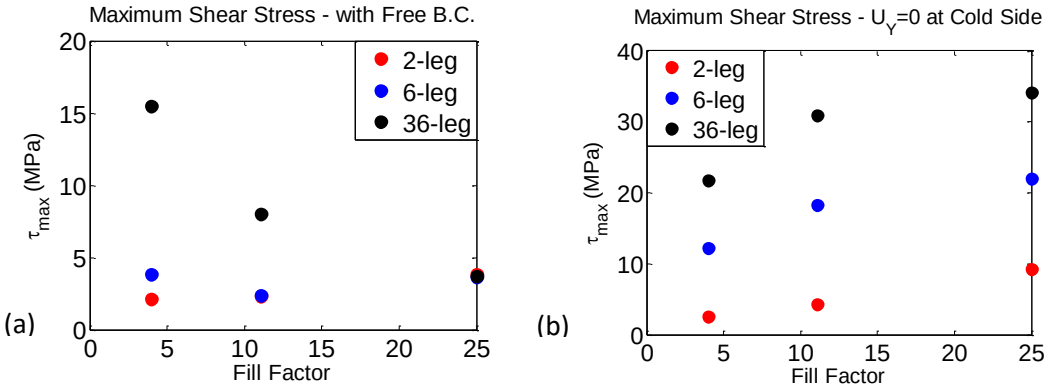


Figure 3.15. Maximum shear stress at the interconnect/solder interface against Fill Factor for three different structures considering different boundary conditions. (a) Free-standing boundary condition; (b) Perpendicular translation is constrained at the bottom interface.

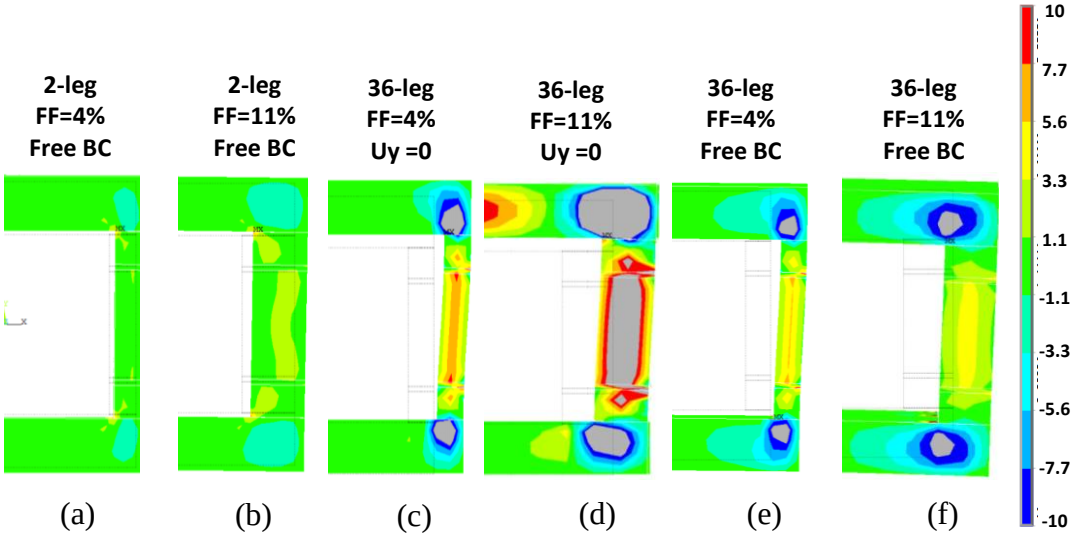


Figure 3.16. Shear stress distribution in TEM structures. Maximum and minimum are capped at ±10MPa. (a) Free-standing 2-leg TEM with FF=4%;(b) Free-standing 2-leg TEM with FF=11%; (c) 36-leg TEM constrained at the cold side with FF=4%; (d) 36-leg TEM constrained at the cold side with FF=11%; (e) Free-standing 36-leg TEM with FF=4%; (c) Free-standing 36-leg TEM with FF=11%.

Shear stress distribution for 36-legs structure under both boundary conditions, and for the 2-leg TEM in free-standing case is demonstrated in Figure 3.16. This figure elucidates why the trends are different for 36-legs and 2-legs in a free standing structure. Figures 3.16a, 3.16b, 3.16e and 3.16f, show that by varying the fill factor from 11% to 4% in a 2- leg design the leg bending decreases (3.16a, 3.16b) while this increases for 36-leg design (3.16e, 3.16f). As a result, the maximum shear stress has an opposite trend in the two cases. Also, Figure 3.16c, and Figure 3.16d, illustrate that when the perpendicular translation is restrained at the cold side, strong stresses is generated on that side. These stresses are larger for the case with larger fill factor and relax in the leg toward the hot side. Therefore, larger fill factor with constrained boundary condition in 36-leg design has larger stress compared

to lower fill factor, which also oppose the trend for the case shown in Figure 3.16e and 3.16f, but follows the analytical predictions.

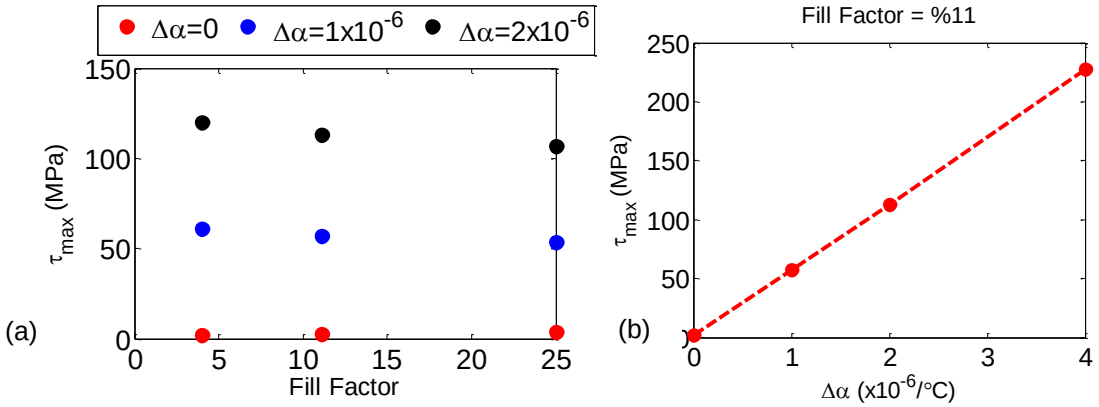


Figure 3.17. Effect of CTE mismatch and Fill Factor on maximum shear stress. (a) Trend of maximum shear stress against fill factor for different CTE mismatches between the solder and its neighboring layers; (b) Maximum shear in TEM vs. CTE mismatch between solder and TE leg material.

In the previous analysis a homogenous coefficient of thermal expansion (CTE) among the materials used in the TEM structure is assumed. While this assumption is not realistic, it gives us an intuition on what is the sole outcome of varying geometrical factors under certain boundary conditions. Under homogenous CTE, only temperature difference between the top and the bottom substrates provoke the maximum shear stress in the structure. However, localized CTE mismatch between the layers could adversely affect the maximum shear stress in the structure and can result in large failure even at small fill factors. Figure 3.17a shows how the trend of maximum shear stress against fill factor for different CTE mismatches between the solder and its neighboring layers. When there is no mismatch, the trend is decreasing and the stress values are low. However, a slight mismatch at high temperature could generate strong stresses as well as changes the trend how stress

scales with the fill factor. Figure 3.17b shows that the maximum shear stress varies linearly with the difference between CTE of the solder and the TE leg material ($\Delta\alpha$), if CTE in all the other layers remains constant.

3.4.Summary

A large temperature difference between the hot and cold substrates in TEMs could result in large local thermo-mechanical stresses and possible mechanical failure. Analytical model presented in this chapter give some initial trends on how geometry can affect the maximum shear stress. Based on these modeling, decreasing fill factor by a factor of 8 in a TEM from 25% to 3%, could result in more than 2x reduction in the maximum shear stress. Finite Element 2D and 3D simulations in ANSYS were carried out to verify the results obtained by the analytical model. Different comparisons are conducted and it is demonstrated that the simple analytical model presented in this work is in good agreement with the results obtained by the Finite Element Method. It is concluded that the maximum interfacial thermally induced shearing stress occurs at the leg's corner and employment of thinner and longer legs could indeed result in a substantial stress relief. One of the main characteristics of an analytical model is that it should be able to distinguish between different parameters and illustrates how variation of each of them would affect the final value of the results. To that end, the significance of this analytical model is that it can be utilized to clarify the effect of different parameters in the model without the need for expensive computation.

FEM is exploited to further investigate the impact of variation in the geometry, boundary conditions, and the coefficient of thermal expansion (CTE) of the materials on the maximum shearing stress in a thermoelectric power generator module (TEM) for high temperature applications. Imposing structural boundary condition on either the hot or cold side of the structure does not change the predicted trend. However, this constraint leads to generation of a stronger shear stress in the contact near the constrained side compared to the free-standing case. The stress will relax toward the unconstrained side. In designing thermoelectric module for high temperature applications, it is not always sufficient to look at the simplest design. Simplifying assumptions in analytical model could limit its range of validity. Depending on the boundary conditions, modules with multiple legs might produce different results compared to the simple two leg design. CTE mismatch between the layers could produce significant local shear stress at the high temperature situation. This local stress should be proportional to $\Delta\alpha\Delta T$. This effect could be mitigated by changing the geometry or choosing material properties that are closely matched.

The numerical examples in this chapter highlight the impact of the fill factor as well as CTE mismatch between different layers. Further work, taking into account the actual manufacturing steps to fabricate a module (e.g. metallization and bonding temperatures) will be necessary to come up with the design of mechanically robust TEM structures.

4. EXPERIMENTAL OBSERVATION OF CURRENT-DEPENDENT PELTIER COEFFCIIENT IN LOW DOPED SEMICONDUCTORS

4.1.Introduction

The Peltier effect is the change in the average energy of electrons when current flows through an interface between dissimilar conductors or semiconductors. When current traverses a junction between two conductors A and B, where in A the average transport energy of electrons is lower than B, electrons will absorb heat from the lattice at the junction to compensate for their lower energy and therefore cool down the junction. This is called Peltier cooling. The amount of Peltier heat carried at the junction is proportional to the applied current by the Peltier coefficient Π. The Peltier coefficient is independent of the electrical current and is related to the Seebeck coefficient by the Kelvin relation, $\Pi = S \times T$. This is a consequence of the Onsager's reciprocity relation. Here S is the Seebeck coefficient, and T is the absolute temperature.

In the case of metallic constrictions, the nonlinear part of the Peltier coefficient may survive at low temperatures and under large electric fields, which could result in thermoelectric cooling [83]. In the nonlinear regime of transport, departures from the Onsager relation between the Seebeck and the Peltier coefficients, as well as from the Wiedemann-Franz law are theoretically predicted [84], [85]. At small systems theses nonlinearities can be observed at relatively small voltages [85], [86]. Bulat had laid out some physical ground work on the nonlinearity of thermoelectric coefficients in early

2000's[87], [88]. Nonlinear responses emerge when transport is strongly nonequilibrium and when electrons are out-of-equilibrium with respect to phonons[34]. Large nonequilibrium at the interface of potential step could be used to increase thermoelectric power generation and refrigeration efficiencies[89]. Zebarjadi, et al. used Monte-Carlo simulations to demonstrate that under strong electric field, electronic temperature starts to exceed the lattice temperature [34]. The non-equilibrium electron heating resulted in the nonlinearity in the Peltier coefficient. Equation (4.1) shows the analytical expression for Peltier coefficient predicted for a non-degenerate semiconductor.

$$\Pi = -\frac{\mu}{e} + \frac{5K_BT}{2e} + \frac{m^*}{2e^3n^2}\left(1 + \frac{10\tau_E}{3\tau_{av}}\right)J^2 \tag{4.1}$$

In this equation , μ is the Fermi energy, e is the electron charge, K_B is the Boltzman constant, T is the absolute temperature, m^* is the effective mass, n is the carrier concentration, τ_E is the energy relaxation time, and J is the current density. τ_{avg} is defined as a characterisitic time which describes how the distribution function relaxes [90]. The nonlinear term in the Peltier coefficient is independent of the temperature, and therefore can survive at low temperatures leading to enhanced thermoelectric cooling [34]. It is also apparent that this coefficient is proportional to the effective mass and inversely proportional to the square of the carrier concentration of a semiconductor. This means in semiconductors with low carrier concentration at large electric fields and current densities, the nonlinearity is more prominent.

Non-equilibrium thermo-electric effects, and in particular, Peltier coefficient has been the subject of few recent studies [91]–[93]. Sadeghian et al. used Monte Carlo simulation to calculate the nonlinear thermoelectric properties of $InAs_{1-x}Sb_x$. They calculated current-

dependent Peltier coefficient, and found the optimum carrier concentration and material properties for a hypothetical thermoelectric module, to maximize cooling efficiency. Muscato and Di Stefano [92] employed a hydrodynamic model to study the equilibrium and off-equilibrium Peltier coefficient in silicon semiconductors. Current-dependent Peltier coefficient in silicon at larger field predicted with qualitatively similar behavior as current-dependent Peltier in InGaAs [34], although the rate of increase was larger than square of the current density. Terasaki et al. [93] investigated two different approaches of nonlinear conduction and photo-doped carriers to benefit from nonlinearity of transport properties in improving thermoelectric power generation. In the former approach, they have studied nonlinear conduction in Mott insulator Ca_2AuO_4. At near room temperatures (300-320K), they measured a factor of 10 decrease in resistivity and 50-100μV/K enhancement in seebeck coefficient due to 100x increase in current density, which is beneficial in improving efficiency of thermoelectric generation. They did not ascribe this enhancement in seebeck coefficient to hot electrons due to strong electron-phonon coupling in Ca_2AuO_4; however, the origin of this enhancement is still under investigation.

In an effort to measure the predicted nonlinear Peltier coefficient in low doped semiconductors and the corresponding cooling, a set of thin film InGaAs microrefrigerators is designed and fabricated. Using detailed thermo-electrical measurements, the current-dependent Peltier coefficient is successfully extracted. The temperature change on top of these devices is obtained using thermoreflectance thermal imaging (TRI) at different current densities. Exploiting TRI results along with four-point probe electrical characterization, it is possible to extract the Peltier coefficient at different current densities

and separate it from non-linear Joule heating in the device. In the following sections we describe the methodology and present the extracted Peltier coefficient. We also discuss how this will affect the amount of the Peltier cooling/heating and in turn the performance of microrefrigerator devices.

4.2.Experimental Methodology

4.2.1.Microrefrigerator Fabrication

A 5μm thick low-doped n-type InGaAs was grown by the molecular beam epitaxy. A highly-doped InGaAs contact layer was grown on top of the active layer for good contact. Devices of various sizes ranging from 10x10μm^2 to 150x150μm^2 were fabricated on the wafer by creating 0.8μm deep square mesas of different sizes using the inductive-coupled plasma etch. A 300nm silicon nitride (Si_3N_4) insulation layer was deposited everywhere, and etched away using the reactive ion etch on top of the mesas for top contact metallization. A side contact is also fabricated for each of the devices to extend the top metal contact to the side with a larger size for easier probing and less probe heat load on top of the cooler. Similarly, a shared ground contact was deposited beside the devices on a much larger area. SEM, top view, as well as cross-sectional view of the device are shown in Figures 51a, 51b, and 51c, respectively.

4.2.2.Electrical Characterization

Four-point probe measurements, using Keithley 2400 source meter, are conducted to obtain the current voltage characteristic (I-V characteristic) of microrefrigerators with different sizes. A typical I-V characteristic for 75×75μm^2 device at 297K and 50K is plotted in

Figure 4.2a. Due to presence of two back to back Schottky diodes, devices act as Metal-Semiconductor-Metal. Under forward and reverse biasing conditions, devices exhibit an almost symmetric characteristic. To study the I-V characteristic, we employed the analysis in Refs. [94]–[98]. The main difficulty arises due to the series connection of two metal-semiconductor-metal structures in the electrical current path, as well as the parasitic and substrate resistances. The first parasitic resistance is at the junction between top metal contact and the thin film, and the second is at the junction (buried junction) between the thin-film and the highly doped substrate. We eliminate some of the parasitics by measuring the voltage of the neighboring devices which follows the equipotential lines from the highly doped substrate. By subtracting the I-V curve at the highly doped substrate (the buried junction) from the total I-V characteristic, the voltage drop across the low doped InGaAs thin film is measured. The carrier concentration of thin film is extracted from the flat band voltage [95] as well as from the temperature dependent measurement of the barrier height [99], and it is estimated to be on the order of $\sim 1.5\times10^{14}$ cm^{-3}.

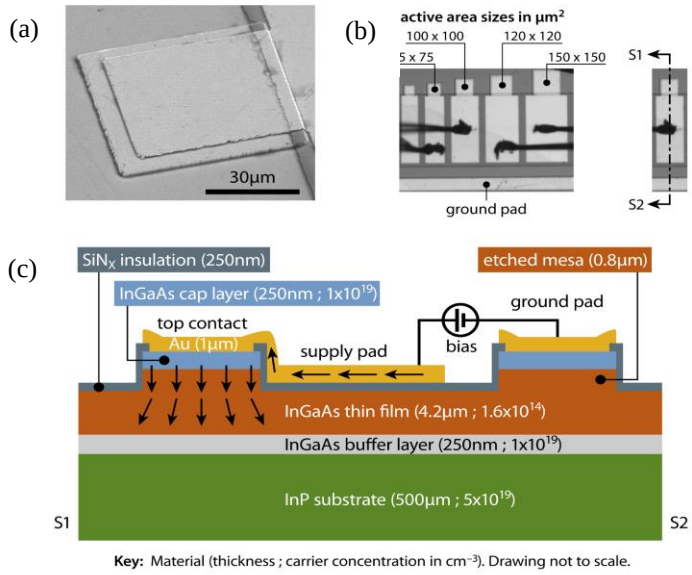


Figure 4.1. Nonlinear Microrefrigerators. (a) Scanning Electron Micrograph of a 75×75μm^2 device. (b) An array of devices. 4 larger size devices as well as the ground pad are wirebonded. Those four devices and the ground contact line are highlighted in the image. (c) A 2D cross section view of different layers in the device along with their dimensions and carrier concentration. InGaAs carrier concentration is derived from I-V characteristic. (The image is not to scale).

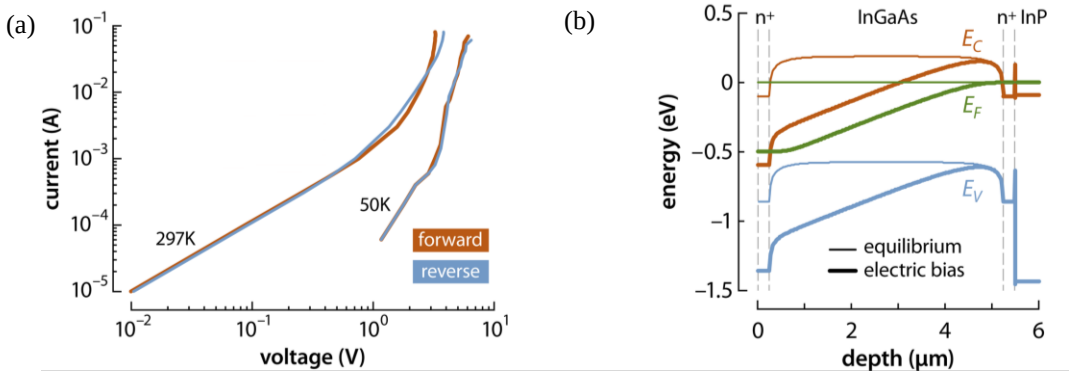


Figure 4.2. I-V characteristic and band diagram of the microrefrigerators. a) I-V characteristic of 75x75μm^2 device at room temperature and 50K. Red curve is the I-V under forward bias (positive current) and blue curve is the I-V under reverse bias (negative current). (b) Band diagram (conduction and valence bands as well as quasi Fermi level). The solid line is the band diagram at equilibrium. The dashed line is the band diagram under bias. Band diagram is plotted using Padre simulator from nanohub [100].

4.2.3.Thermal Characterization

The thermoreflectance imaging microscopy is utilized for thermal characterization of microrefrigerators. In this technique, to measure the temperature of the device under test, the change in the reflectivity of the surface due to temperature change is calibrated [29], [101]. The technique is not only beneficial in quantifying the surface temperature, with submicron spatial and 10mK temperature resolution,[102] but also it can be used to visualize the current uniformity in active devices [103], [104]. The experiments are conducted in the high vacuum environment (~1E-6 Torr) to prevent the heat loss through the air convection as well as water condensation at lower temperatures. For room temperature measurements, the temperature is set to 297K. Positive and negative currents are applied in forward and reverse biases. The device under test, as well as the temperature profiles under forward and reverse biasing are depicted in Figures 4.3a, 4.3b and 4.3c. The temperature profiles are calibrated for Gold, with coefficient of thermoreflectance or C_{TR} = -2.3×10^{-4} /K, and for InGaAs, with $C_{TR}=1.3\times10^{-4}$ /K. We masked the Si_3N_4 region in the profiles. The applied electrical currents are 71.2mA and -74mA in forward and reverse directions, respectively. The cross sections of the temperature profiles under both polarities are plotted in Figures 4.3c and 4.3d. Each data point is obtained by averaging over a box of $7\times45\mu m^2$ shown in Figure 4.3a. This box is swept along A-A' direction. By changing the bias polarity, one can separate the reversible Peltier and the irreversible Joule component of temperature change at the top interface. The Peltier signal depends on the direction of current and changes signs by switching the polarity. This means the Peltier heating and cooling occurs under forward and reverse biasing, respectively. On the other

hand, Joule heating is proportional to current multiplied by voltage and will have the same sign in both polarities.

$$\Delta T_{Fwd} = \Delta T_{Peltier} + \Delta T_{Joule\_Forward} = aI + b_1 I V_{Fwd} \quad (4.2)$$

$$\Delta T_{Rev} = -\Delta T_{Peltier} + \Delta T_{Joule\_Reverse} = -aI + b_2 I V_{Rev} \quad (4.3)$$

Here, ΔT_{Fwd} and ΔT_{Rev} are the amount of change in temperature on the top surface in forward and reverse directions. $\Delta T_{Peltier}$ and ΔT_{Joule} are the Peltier and Joule contributions to the change in the surface temperature. I and V are the current and the voltage (forward and reverse), and a, b_1, and b_2 are the Peltier and Joule terms coefficients. In a linear Ohmic device, Seebeck voltage is relatively small compared to electrical bias. V_{Fwd}, V_{Rev}, b_1 and b_2, and hence forward and reverse Joule heating do not change by switching the polarity. Consequently, the Peltier and the Joule effects may easily be separated. However, as it can be seen later on, due to asymmetry of the Joule heating in the nonlinear microrefrigerators, under forward and reverse polarities, $\Delta T_{Joule\_Forward}$ and $\Delta T_{Joule\_Reverse}$ are not the same. The small asymmetry in the forward and reverse I-V characteristics is indicative of asymmetric Joule heating under either biasing. Additionally, the stronger non-uniformity in the temperature profile under forward bias compared to the reverse bias confirms the argument on the location dependent Joule heating in the devices. Under forward bias most of the Joule heating occurs near the top interface while under reverse bias most of the Joule heating occurs near the bottom interface. Due to the current spreading, the cross sectional area seen by the current near the buried interface is larger which cause the heat generated near that interface in reverse bias to be less concentrated than the forward bias. Consequently, a more uniform temperature profile is

obtained under reverse bias. Moreover, the tail of the temperature cross sections in Figure 3d. has a larger value (~4K) in the reverse bias than in the forward bias. This signifies the region near the buried junction is hotter under the reverse bias. This is expected as both the Joule and the Peltier heating under the reverse bias take place near and at the buried interface, respectively, while the Peltier cooling in forward bias is at the same interface. On the other hand, it is clear from the figures that the regions near the top interface is colder under the reverse bias. This is also expected, as the Joule and Peltier heating under the forward bias occur near and at the top n^+-n interface, respectively, while the Peltier cooling under the reverse bias is at the same interface. ANSYS [66] FEM simulations are used to verify this explanation. In section III, we separated the Peltier and Joule components in the forward and the reverse biasing. The separated Peltier and Joule heating are input to the ANSYS model to calculate the temperature profiles at the top surface. The red and blue dashed lines in Figure 4.3d are ANSYS results for the forward and reverse biasing, respectively. It should be pointed out that in ANSYS modeling the Joule heating source was near the top junction at forward polarity to accurately obtain the non-uniform temperature profile, while in the reverse polarity the Joule heating source had to be further away from the top interface and closer to the buried junction. This further validates the argument for location dependent Joule heating.

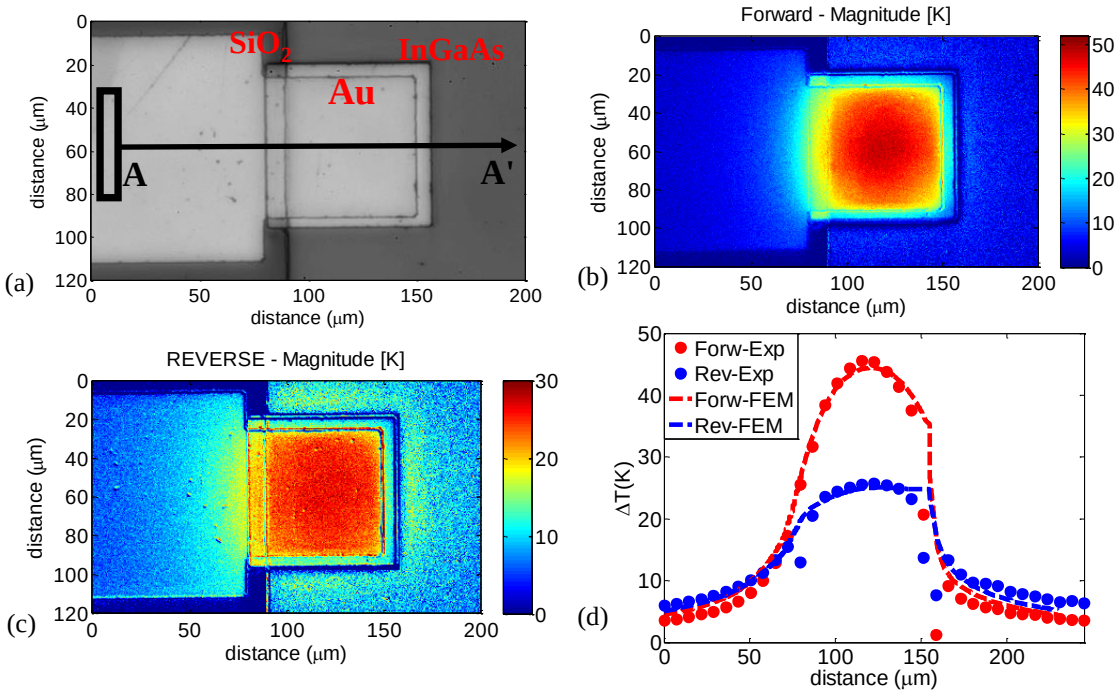


Figure 4.3. Thermal Imaging of Microrefrigerators. (a) A CCD image of the device under test. Au contact, SiO_2 insulator and InGaAs film regions are specified in the figure. (b) Temperature profile under forward bias. I=71mA, V=3.76volt. C_{TR} is -2.3×10^{-4} for Au, and 1.3×10^{-4} for InGaAs. SiO_2 is masked in the profile. (c) Temperature profile under reverse bias. I= -74mA, V= -4.3volt. (d) Cross section along A-A'. The cross section is obtained by sweeping a box average ($7\times45\mu m^2$) along A-A'. The dashed lines are the ANSYS simulation results considering heat sources near the top (Forward) and buried (Reverse) junctions. The amount of heat is determined based the extracted Peltier and Joule heat components from the experimental data.

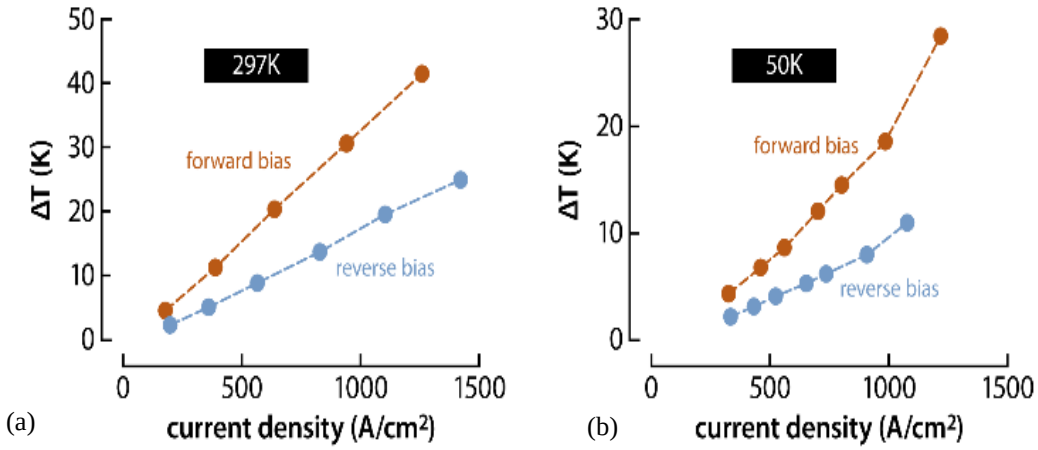
4.2.4.Results and discussions

We developed a hybrid analytical-numerical scheme based on the full heat balance equation as well as the equivalent thermal network of the devices to separate the Peltier and Joule components of temperature from which the Peltier coefficient is extracted. We constructed a 1D thermal network with the thin-film as well as substrate thermal resistances. The 3D heat spreading is included through analytical equations verified by

ANSYS. The details of this model is described in appendix B.1. The average temperature change on the top surface of the 75x75μm^2 microcooler in forward and reverse biases, for 297K and 50K ambient temperatures, are plotted against current density in Figures 4.4a and 4.4b, respectively. By increasing the current, the average temperature in the forward bias surges to larger values than in the reverse bias. One should note that the error bars in these figures are not indicative of the error in the measurement, rather they are standard deviations in the distribution of temperature on the top surface. Under forward and reverse biases, Peltier heating and cooling occurs at the top and the bottom interfaces, respectively. Additionally, as described in section 4.2.3, Joule heating under forward and reverse polarities is generated near the top and buried junctions, respectively. Therefore, both Peltier and Joule effects contribute to the large temperature difference between forward and reverse polarities.

The extracted Peltier coefficient at room temperature is plotted against current density (A/cm^2) in Figure 4.5a. The blue, red, and green dots are extracted Peltier coefficients for 75×75μm^2, 100×100μm^2, and 120×120μm^2, respectively. We compared our results with theoretical predictions by the analytical equation (4.1). The energy relaxation time (τ_E) and distribution function relaxation time (τ_{avg}) are obtained from the Monte Carlo model developed elesewhere [34], [91]. As discussed in section 4.2.2, the estimated carrier concentration of the InGaAs layer is about ~1.5×10^{14} (cm^{-3}), for which we plotted the analytically predicted Peltier coefficient. The experimental and analytical results are in good agreement until about 500-600A/cm^2. Beyond these values, the rate of increase in the experimentally measured Peltier coefficient is lower compared to the analytical results.

This can be explained by the barrier lowering at the emitter junction. An approximate band diagram is shown in the inset for Figure 4.5a. At small voltages, the barrier at the emitter prevents the low energy electrons from contributing to the transport. By increasing the voltage and in turn the electric field and the current density, more hot electrons above the Fermi level contribute to the transport, hence the Peltier coefficient increases. As we apply larger voltages, the voltage on the cathode starts to lower the barrier on the emttier junction and causes the electrons below the Fermi level also contribute to transport. This acts as a competing effect to reduce the Peltier coefficient and as a result the rate of increase in the Peltier coefficient decreases. This is similar to drain-induced-barrier-lowering in MOSFETs [105]. The Peltier coefficient is also extracted at cryogenic temperatures. Figure 5b shows the Peltier coefficient at 30K, 50K and 70K. The results compared with the theoretical predictions at 50K (The difference between the results at 30K, 50K, and 70K is negligible). The extracted Peltier coefficient follows the same trends as the theoretical predictions up to current densities of 1200-1300A/cm^2.



(a) (b)

Figure 4.4. Average temperature change at the top surface of 75×75μm2 device under forward (red) and reverse (blue) polarities vs current density (A/cm2). (a) At 297K; (b) At 50K.

The amount of Peltier induced cooling at room temperature is plotted in Figure 4.5c and compared with the expected nonlinear and linear cooling. The approximate linear cooling is obtained with a linear InGaAs microrefrigerator with the same material properties and the linear (current independent) Peltier coefficient. As it can be seen in Figure 4.5c, the cooling is almost increased by more than a factor of two at room temperature compared to the linear device, and as theoretically predicted, it could have been increased by a factor of five. Although, the measured nonlinear Peltier cooling is significant, it is not sufficient to overcome the excessive Joule heating in the junctions, the thin film, and the substrate, that is due to large current densities and the small thermal conductivity of the InP substrate (~73W/m-K). The amount of Peltier cooling at cryogenic temperatures is plotted in Figure 4.5d and is compared to the nonlinear and linear predictions. The maximum cooling is about 20x larger than the theoretical prediction for the linear device at the same current

density. Again, no net cooling is observed. This is despite the fact that at those temperatures due to significant increase in substrate (InP) thermal conductivity, ~600W/m-K, ~1000W/m-K and 2000W/m-K at 70K, 50K and 30K, respectively [106], the Joule heating from the substrate is almost negligible. It can be inferred from Figure 4.2a that the main reason for the net heating is the Joule heating from the thin film. As it is evident from this figure, at 50K there is almost twice as much voltage needed to reach the current densities at which the nonlinearity takes place.

Therefore, although current dependency of Peltier coefficient leads to a significant increase in the Peltier cooling, the extreme Joule heating from the substrate at room temperature and from the thin film at cryogenic temperatures prevented net cooling in the device (See supplementary materials). A substrate transfer to a metallic substrate can reduce the effect of Joule heating from the substrate at the room temperature and decrease the total Joule heating by a factor of 2.5x.

Transient TR thermal microscopy[27], [31] was used to further investigate the Peltier effect in nonlinear microrefrigerators. We applied a 1μs current pulse with %1 duty cycle under both forward and reverse polarities to a $50\times50\mu m^2$ microrefrigerator. The results are shown in Figure 4.6. Figures 4.6a and b show the transient response of the microrefrigerator under forward and reverse polarities. It is apparent from the figures that Peltier cooling occurs under reverse biasing. As it is expected, since Peltier takes place at the top junction, it manifests itself prior to Joule heating. Moreover, it should be noted that the maximum temperature change at the top surface under forward bias occurs about 3μs before that of reverse bias.

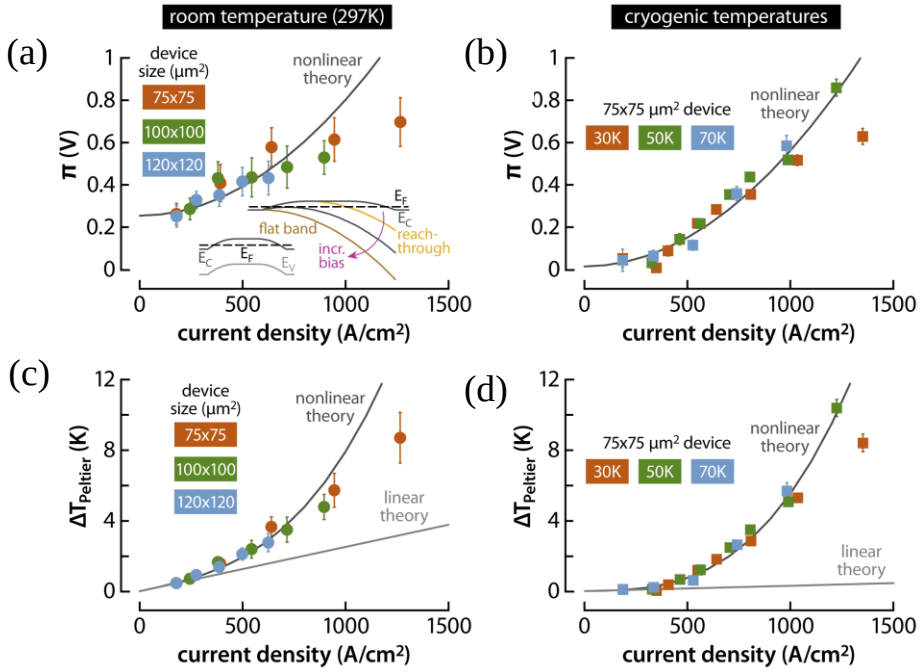


Figure 4.5. Nonlinear Peltier coefficient and cooling at room and cryogenic temperatures. Peltier coefficient against current density: (a) for three different device sizes. Black line is analytical model prediction with carrier concentration 1.5×10^{14}cm^{-3}. Inset: Band diagram under equilibrium as well approximate band diagram that shows barrier lowering at the emitter junction; and (b) At 30K, 50K, and 70K for 75×75μm^2. Black line is the analytical prediction for 50K. (c) Experimental Peltier cooling against calculated nonlinear and liner Peltier (at 297 K). (d) Experimental Peltier cooling against calculated nonlinear and liner Peltier at cryogenic temperatures.

This is an indicition of location dependent Joule heating. Under forward bias Joule heating is closer to the top surface and causes the temperature at top reaches its maxmium faster than reverse bias. Normalized temperature is plotted against time in Figure 4.6c for both forward and reverse polarities. The decaying part of the curve is plotted in logarithm scale in Figure 4.6d. Assuming Fourier heat equation for the decaying curve, the response has to be proportional to $exp(\frac{\pi^2\alpha t}{4l^2})$ with the time constant $\tau = 4l^2/\pi^2\alpha$, in which α is the thermal

diffusivity of thin film and l is the thickness[107]. Assuming $\alpha = 2.27\times10^{-6}$ m^2/s for InGaAs [106], and using the slopes shown in Figure 4.6d, l was obtained to be 4.8µm under forward bias, which is about the same as thin film thickness, and about 9.3µm under reverse bias. This demonstrates that back flow of Joule heating under forward and reverse bias was not the same as we explained in in the text and appendices.

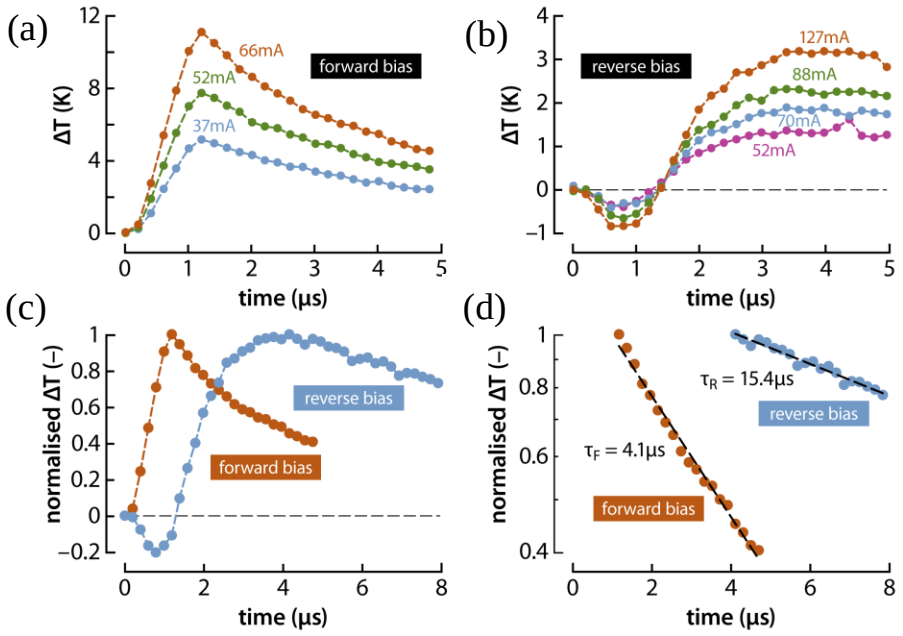


Figure 4.6. Transient response of a nonlinear microrefrigerator. $50\times50\mu m^2$ Microrefrigerator response to a 1µs current pulse under (a) Forward, and (b) Reverse bias. (c) Evolution of normalized ΔT over time at both polarities. (d) The decay of normalized ΔT against time. The dashed lines show the fitting against time. τ_F and τ_R represent the time constants under forward and reverse polarities, respectively.

4.3.Summary

The first room and cryogenic temperature observation of the current-dependent Peltier coefficient is presented. About 2.5x increase in the Peltier coefficient at room temperature and 20x increase at cryogenic temperatures is observed. Although, the amount of Peltier-

induced cooling obtained is significantly larger than its counterpart linear devices, the net steady-state cooling is not observed due to significant Joule heating from substrate at room temperature and from the contacts and the thin film at cryogenic temperatures. High bias barrier lowering was identified as the main impediment to benefit from huge Peltier coefficient at very large current densities. Transient thermal characterization does show net cooling by ~1C at room temperature at a time of about 1μs after a current pulse of 5080 A/cm^2. Further optimization of the material is necessary to benefit from nonlinear Peltier and obtain significant net cooling. The presence of metal/low-doped semiconductor contact is needed to observe the nonequilibrium and nonlinearity. In order to avoid the space-charge regions and barrier lowering effects at high biases, it may be possible to benefit from delta-doping in the barrier which will cause the applied electric field to be shielded from the injecting electrode. Similar structures have been used in ballistic electron emission microscopy (BEEM) for fine control of the energy of the ballistic electrons.

5. SUBDIFFRACTION LIMIT THERMAL IMAGING

5.1.Introduction

High-speed switching power transistors, such as high electron mobility transistors (HEMT), working in frequencies on the order of tens of gigahertz have sophisticated designs with cross-sectional T-shape gate structure with very small gaps and long straight electrodes, to enable high power densities for wireless mobile applications. However, increased Joule heating at large current densities will cause a rise in the junction temperature, which in turn reduces carrier mobility, and thermal conductivity. This can also result in microscale thermomechanical stresses. These will have adverse impacts on the performance, long term reliability and lifetime of HEMT devices [108]–[111].

One of the challenges for modeling and parametrizing transistor performance is to accurately measure the temporal temperature response of the device under operating conditions. Thermoreflectance thermal imaging (TRI) as a noninvasive optical technique suitable for 2D mapping of temperature field in active semiconductor devices such as in HEMTs [33]. Despites its higher resolution compared to other available optical techniques such as IR and micro-Raman, TRI can still be limited by the optical diffraction for thermal measurement of the devices with sub-diffraction features. In this chapter we systematically investigate the effect of optical diffraction on the TRI measurement for small heater lines ranging from 100-1000 nm widths. We address the impacts of optical diffraction on the

accuracy of the technique and how one can compensate for it. We introduce an analytical model that can be used along with numerical modeling to accurately predict the temperature profiles of sub-diffraction devices. Further, we show how using image processing reconstruction techniques, we are able to treat the ill-posed inverse problem and accurately reconstruct the temperature profiles of devices with sub-diffraction features. In Section 5.2, we discuss the experimental technique, numerical and analytical thermal modeling. In Section 5.3, we describe the thermal images with apparent temperature profile. Section 5.4 summarizes the reconstruction algorithm and more accurate thermal profiles. We conclude in Section 5.5.

5.2.Methodologies

5.2.1.Thermoreflectance Imaging Microscopy

Thermoreflectance (TR) imaging microscopy is already introduced in chapter 4. In the following sections, we use TR thermal imaging microscopy to measure the temperature profile of heater lines ranging in width from 100nm to 10μm, and analyzed the effect of the optical diffraction on the measured temperature profile.

5.2.2.Nano-heater-line fabrication

A set of heater lines with different widths ranging from 100 nm to 10 μm, which includes some sub-diffraction-limit widths, were fabricated. After removing the native oxide on the film with 1 minute dilute HF solution, a 20 nm Al_2O_3 insulation layer was deposited using the atomic layer deposition technique at 200^0C on $In_{0.53}Ga_{0.47}As$ (5 μm) / $In_{0.52}Al_{0.48}As$ (100 nm) / InP (500 μm) sample (grown by molecular beam epitaxy) followed by rapid thermal annealing at 450 ^{0}C for 30s. The samples were then processed using electron beam

lithography (EBL), metallization and then lift-off to obtain the Au (~85nm) / Ti (5nm) heater lines of different widths. The InP substrate is sufficiently thick, and therefore, semi-infinite medium is assumed. Aspect ratio of each device, i.e. ratio of length to width, was fixed to 20. Four large contact pads, each 80×80 μm^2, were fabricated for each heater line, so that the samples can be probed easily and also the thermal measurement can be further confirmed using the heater resistance. Two similar heater lines were placed in parallel next to each other with a certain distance (gap), which is varied from 200 nm to 500 µm (semi-infinite). The effect of the gap size on the temperature profile is a topic for a subsequent study and will not be discussed in this chapter. TR imaging is done using a LED with 530nm wavelength to illuminate the device and a ~250K pixel single photon counting industrial class CCD imager. Optical images for 200nm and 1000 nm devices are shown in Figure 5.1a and 5.1b, respectively.

Four-point probe technique was used to measure the resistance of the gold metal lines at different ambient temperatures, and subsequently the temperature dependent electrical resistivity and thermal conductivity of the metal line are determined. These values were used for the subsequent finite element modeling of the heater lines. Additionally, electrical IV measurements of the four probe structure is used to extract the average temperature on top of the heater line at different current values. This is explained in chapter 6.

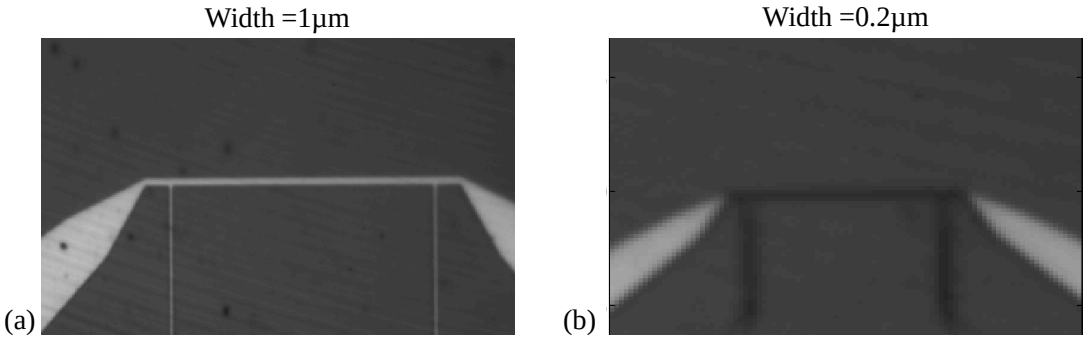


Figure 5.1. Optical CCD images of nanoheaters lines. (a) a 1μm wide Au heater line, and (b) a 200nm heater line, obtained with 100x objective lens and 250K pixel CCD camera. The numerical aperture of the 100x objective lens is 0.75, which suggest a resolution limit of about 353nm for LED source light of 530nm.

5.2.3.Analytical Modeling

The Rayleigh resolution criterion states that two point objects are resolved in distance *D* (radius from the point source) given by [55]:

$$D = 1.22\lambda N \tag{5.2}$$

where, *N* is the f-number of the objective and λ is the wavelength of the light reflected from the surface. *N* is approximately (*1/2N.A.*), where *N.A.* is the numerical aperture of the lens. Equation (5.2) is commonly known as the diffraction limit. We are measuring a device width that is smaller than the distance, *D*, a sub-diffraction-limit feature. Diffraction limit is determined by the emitted or reflected light from a point source since higher spatial frequency light from the source is filtered by the optical system. Airy disk determines the minimum distance from a point source at which the imaged light intensity goes to zero in the Bessel function profile as a function of distance [112]. This intensity profile can be approximately fit with a Gaussian profile with full-width-half-maximum 2.44 times the Airy disk diameter, which is convenient for further mathematical formulation (see Figure

5.2). The light intensity as a function of distance from a point source is described as $E = E_0 f(r)$ with,

$$f(r) = \left(\frac{2J_1(\pi r / \lambda N)}{\pi r / \lambda N} \right)^2 \cong \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(2.44 r / \lambda N)^2}{2} \right) \tag{5.3}$$

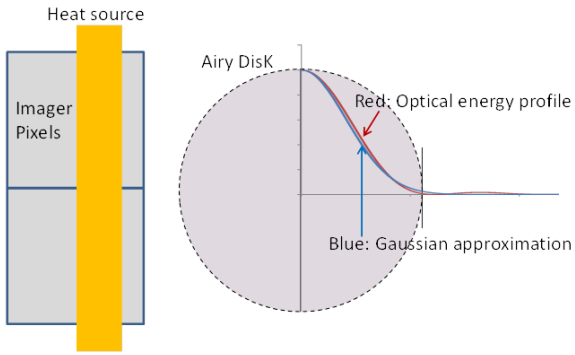


Figure 5.2. Schematic of heat source 100 nm wide, imager pixel 300×300 nm2, Airy disk ~ 1.3 µm (if N=1). It is assumed that the heater line is parallel to the line of pixels but not aligned to the center of pixels.

5.2.4. Finite element numerical modeling

A electrothermal finite element (FE) model is constructed in ANSYS Parametric Design Language (APDL) for each of the devices. The material properties are input to the model based on the independent characterization as discussed in section 5.2.2. Thermal conductivity of each layer is measured using 3ω and TDTR techniques. The meshed structure of the modeled 200 nm heater line in ANSYS along with its temperature profile is shown in Figure 5.3.

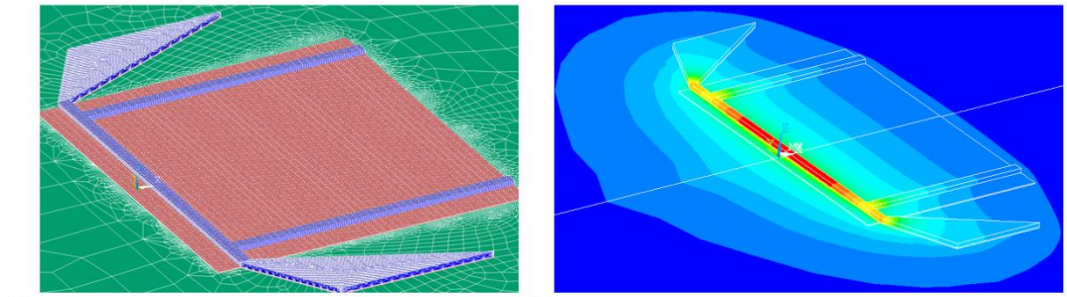


Figure 5.3. ANSYS modeling of nanoheaters lines. (a) Meshed structure of the modeled 200 nm heater line in ANSYS. The length of the heater line is 8 μm (Aspect ratio = 40). (b) associated temperature profile for the heater line shown in part a.

The numerical results from ANSYS were compared with the experimental TR images. Based on careful calibration, the coefficient of thermoreflectance for Gold for 100x objective for 1μm device is ~ -2.2×10^{-4}, and the coefficient of thermoreflectance for the substrate area outside the gold lines is found to be ~ $+2\times10^{-4}$. Experimental results are normalized with the gold coefficient only, therefore the ANSYS results on the substrate is normalized by the ratio of the substrate to gold C_{TR} for comparison purposes. Figure 5.4, shows the experimental results for two case studies on a 1μm device with two different objective lenses, 100x and 10x. The numerical aperture for 100x and 10x objective lenses are 0.75, and 0.2, which corresponds to Rayleigh radius of 353nm and 1330nm, respectively. This means that the 1μm heater line is well above diffraction at 100x, while is below diffraction at 10x. It is worth noting that each pixel size at 100x is about 160nm, while it is 1.43μm at 10x. This means a 1μm heater line at 100x contain 7 pixels, while a fraction of a pixel is filled at 10x. This is visible in the optical CCD images shown in Figure 5.4a and b. Figure 5.4c and d shows the temperature map obtained for each case for 1μm device. We used a constant coefficient of thermoreflectance in both cases. Cross sections

of temperature profile along the vertical axis is compared in Figure 5.4e. We also plotted the cross section obtained from FEM.

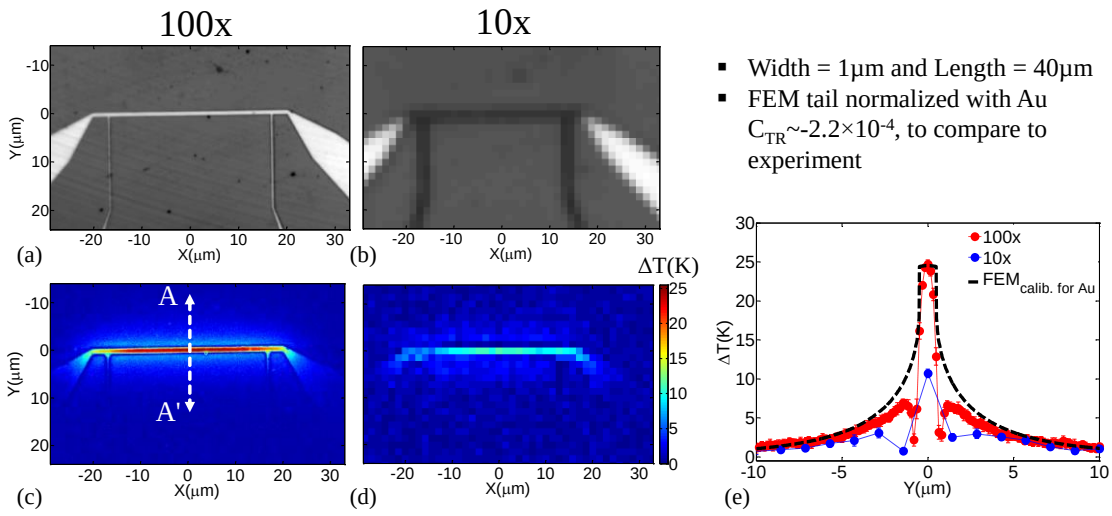


Figure 5.4. Optical and thermal image of 1μm heater line with different objective lenses. CCD images under (a) 100x, and (b) 10x, objective lenses. Corresponding TR thermal images for the 1μm heater line under (c) 100x, and (d) 10x. (e) A comparison between the cross section of the temperature profile along A-A' for FEM, TRI at 100x and 10x. The profiles and FEM are calibrated and normalized for gold C_{TR}, and the absolute value of corresponding temperature profile is plotted. It is evident that due to diffraction limit the apparent temperature measured with the 10x is underestimated.

From these figures two major differences are evident:

1. In the 100x image, where the heater line width is larger than the diffraction limit, the only differences between the ANSYS and experimental results are found at the edges of the devices. The edge effects arise due to the fact that the coefficient of thermoreflectance of gold and the substrate are opposite in sign. Of course due to nature of diffraction function the shape of the temperature profile at the top is less sharp compared to modeling. Within few microns outside the edge of the heater line there are some differences in the tail that is explained in chapter 6 due to non-diffusive heat transport.

2. In the 10x image, where the heater line width is smaller than the diffraction limit, there are not only edge effects but also the value of apparent temperature on top of the device changes. This means a much smaller coefficient of thermoreflectance (C_{TR}) is needed for the devices that are diffraction limited to obtain the true temperature value.

Before continuing with the diffraction effect on thermal images, it is important to answer why we observe thermal images despite the fact that the device size is below diffraction? The answer is believed to be similar to the idea of phase shift mask in photolithography. In photolithography, in order to discern two objects that are closer than diffraction limit, a 180º phase shift mask is used to inverse the electric field under the aperture for one of the objects which in turn results in separating the intensity signal from the two objects and discern them. Here we have a metal on top of a semiconductor that have opposite sign C_{TR} at the LED wavelengths we are using. This difference in C_{TR} in the thermal domain causes different sign temperatures which in turn creates zero crossing near the edge of the metal and therefore the signal from metal and the substrate can be distinguished.

Since the diffraction function works similar to a blurring point spread function in image processing, we have to able to obtain the 10x results from 100x results by simply convolving the latter one with the diffraction function, that is explained in section 5.2.3 and equation 5.3. To do that, we normalized the 100x temperature profile with gold coefficient everywhere (Figure 5.5 a1), convolved the results with the Gaussian intensity filter simulating the optical diffraction function (Figure 5.5 a2), taking into account the LED light wavelength and numerical aperture of the objective lens, and obtained the blurred temperature map (Figure 5.5 a3). A comparison between the results are shown in

Figure 5.5b. It is apparent that the amplitude of temperature at the top for the blurred map matches the results obtained at 10x for which the thermal image is shown in Figure 5.4d.

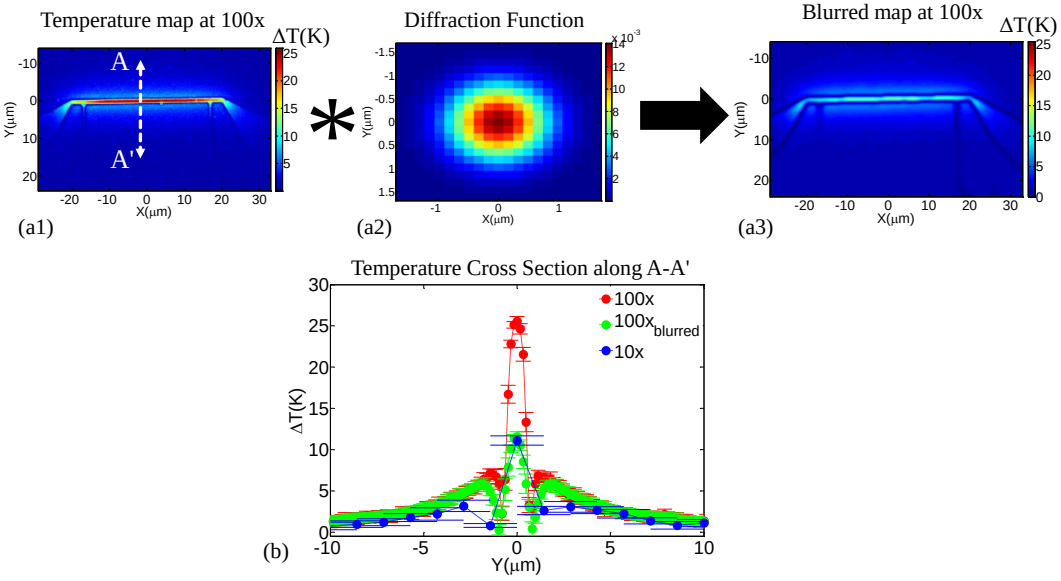


Figure 5.5. Effect of diffraction function on thermal imaging. (a) experimental thermal image of a 1μm heater line obtained with 100x objective lens (a1), must be convolved with the diffraction function (a2) to explain the apparent measured temperature underestimation with a 10x objective lens (a3). Figure 5.5d shows the corresponding experimental results. The resolution limit at 10x and 100x are 0.35μm and 1.33μm, respectively. (b) A temperature cross section comparison along A-A' shows that the blurring by diffraction function accurately matches thermoreflectance measurements with lower numerical aperture lens below diffraction limit.

Therefore, it is evident that the diffraction does play a significant role in the temperature profile obtained from TR imaging. Fortunately, for these heater lines we are able to extract the average temperature and C_{TR} independently by electrical resistivity measurement (see chapter 6). However, in reality it is significantly cumbersome to calibrate for each individual device that is below diffraction and sometimes, such as in nanoscale transistors, is not always possible to have independent electrical measurements for features inside the transistor. In practice we can calibrate the thermoreflectance coefficient using larger areas

on the materials or devices. It will be quite beneficial if the calibration could then be extended for small size devices that are below diffraction limit.

In order to include the diffraction effect in our (forward) modeling, we take the temperature profile determined by ANSYS. Then we generate the thermal image by convolving with the appropriate optical diffraction function. Here is a summary of the steps to obtain the final temperature profile:

1. Obtain the temperature profile from the ANSYS simulation.
2. Normalize this temperature profile using the calibrated value for the coefficient of thermoreflectance for gold.
3. The new temperature profile is then blurred (convolved) with the approximate optical diffraction function. This should produce a temperature profile that is comparable to the experimental results.

In the next section we discuss the results that are obtained from steps 1 through 3, and compare them with experimental temperature profile.

5.3.Results and discussions

The procedure discussed in the previous section is implemented for different devices. Figure 5.6 shows the results for a 1μm device. Figure 5.6a shows a cross section of the Gaussian intensity filter simulating the optical diffraction function at 100x (red), and 10x (blue). N.A for 10x and 100x are 0.2 and 0.75, respectively. It is clear that due to smaller numerical aperture, the Gaussian filter is wider for 10x and therefore resolution is lower (1.33μm compared to 0.35μm). Figure 5.6b and c demonstrate that by convolving the 100x

and 10x filters with FEM results, the blurred response reproduces the experimental results very well.

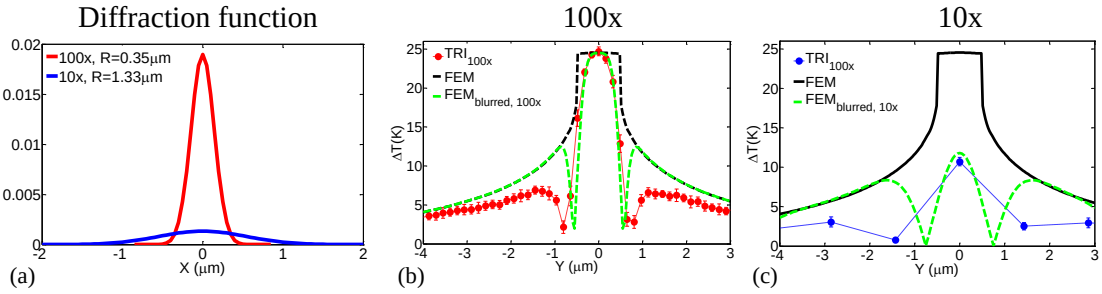


Figure 5.6. Comparison between TR Imaging, FEM and Image Blurring Cross sections. (a) Cross sections of the Gaussian intensity filter simulating the optical diffraction function at 100x (red), and 10x (blue). R is the Rayleigh criterion that set the resolution for the imaging system. Comparison between TR, FEM, and blurred FEM results' cross sections at (b) 100x, and (c) 10x. It is apparent from Figures b and c that as the device becomes smaller than diffraction limit the apparent C_{TR} changes and the temperature values are underestimated. Our blurring model accurately captures this sub-diffraction effect.

Results for a 200 nm heater line is shown in Figure 5.7. First, we obtained the ANSYS temperature profile using the material properties extracted from experiments (Details in chapter 6). This is shown in Figure 5.7a. We then filtered the temperature profile following steps 1 to 3 described above, and compared the resulting temperature profile TRI measurements in Figures 5.7b. It is apparent that the apparent peak temperature has changed by almost 4x due to the diffraction limit. This indeed matches independent electrical calibration of these devices that shows a with $C_{TR} \sim -0.59\times10^{-4}$ for gold at 200nm, one can reproduce the maximum temperature. The typical value for gold at this wavelength (530nm) is about -2.2×10^{-4}. A cross section comparison along $y = 0$ is shown in Figure 5.7c. At the edges, our modeling predicts a dip for the temperature similar to experimental results, which mainly arises due to the change of sign in the C_{TR} at the metal/semiconductor

interface. The main difference is observed within few microns outside the heater lines. This is mainly due to small size of heat source and non-diffusive heat transport that is further explained in chapter 6. Plasmonic resonances could also be a contributor to the edge effect which is subject of a future study. Results for a 200 nm heater line is shown in Figure 5.7. First, we obtained the ANSYS temperature profile using the material properties extracted from experiments (details in chapter 6). This is shown in Figure 5.7a. We then filtered the temperature profile following steps 1 to 3 described above, and compared the resulting temperature profile TRI measurements in Figures 5.7b. It is apparent that the apparent peak temperature has changed by almost 4x due to the diffraction limit. This indeed matches independent electrical calibration of these devices that shows a with $C_{TR} \sim -0.59\times10^{-4}$ for gold at 200nm, one can reproduce the maximum temperature. The typical value for gold at this wavelength (530nm) is about -2.2×10^{-4}. A cross section comparison along $y = 0$ is shown in Figure 5.7c. At the edges, our modeling predicts a dip for the temperature similar to experimental results, which mainly arises due to the change of sign in the C_{TR} at the metal/semiconductor interface. The main difference is observed within few microns outside the heater lines. This is mainly due to small size of heat source and non-diffusive heat transport that is further explained in chapter 6. Plasmonic resonances could also be a contributor to the edge effect which is subject of a future study.

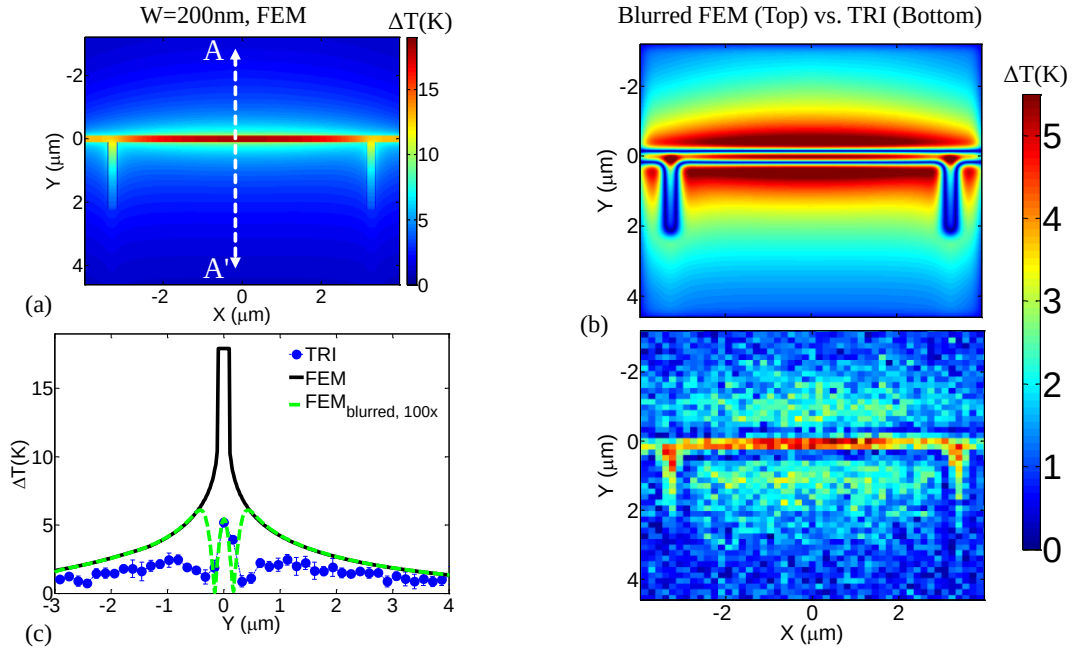


Figure 5.7. Comparison between TR Imaging, and Blurred FEM Temperature profile. (a) ANSYS temperature profile for the 200nm heater line. (b) Filtered result (Top) using a Gaussian function with coefficient 2.44 within an Airy disk with radius of 353 nm. The experimental TR result is also shown in the bottom figure. TRI is done with green LED light with a peak wavelength of 530 nm, and using an optical system with 100x objective lens with N.A. = 0.75. (c) Comparison between the cross sections of temperature profile for the three images, shows excellent agreement at the top of metal line between modeling and experimental results. The difference on the tail of distribution is due to the non-diffusive heat transport which is described in chapter 6.

5.4.Temperature map reconstruction

In practice, it is not possible to measure independently the local temperature for e.g. nanoscale transistors [53], [113], [114] or HAMR devices [115], and forward modeling is not always doable. Only a thermal image is available, and we have knowledge about the imaging system that guide us to obtain the diffraction function. Therefore, an optical reconstruction algorithm is needed to obtain the true temperature profile from the measured apparent temperature map when the image is blurred due to diffraction. Here we will

describe a maximum-a-posteriori (MAP) technique to solve this inverse problem. Assuming *g* is the measured apparent temperature and *f* is the true temperature vectors, we can write the following:

$$g = \boldsymbol{H}f + n \tag{5.4}$$

Here **H** is matrix representation of the blurring kernel (diffraction function), and *n* is a vector representing the noise adulterating the measurement (or modeling). Equation 5.4 is written in lexicographical order. Here we would like to extract *f* from the measured *g*, known **H** and estimated *n*. To solve this problem, a simple inverse process through maximum likelihood (ML) estimation that minimizes the distance (norm 2) of the *g* and **H***f* does not work. This is because a small amount of noise could result in an unstable solution [116]. Instead, in a MAP framework a prior model assumed for the unknown image is used in a regularization term added to the ML estimate to produce the cost function to be minimized. For the prior model we use a non-Gaussian Markov Random Field (MRF) and need to minimize the following cost function:

$$\hat{f} = \underset{f}{Argmin}\{\frac{1}{2\sigma_w^2}||g - \mathbf{H}f||^2 + \frac{1}{p\sigma_x^p}\sum_{\{i,j\}\in C} S_{i,j}|f_i - f_j|^p\} \tag{5.5}$$

We chose p=1 because by incorporating an L1 norm we will be able to better restore and reconstruct the sharp step edges that are expected in thermal problems with rectangular and square heat sources. In our problem we use *S*, [1/12, 1/6, 1/12; 1/6, 0, 1/6; 1/12, 1/6, 1/12], in an 8-point neighborhood systems determined by *i* and *j* pixels in image *f*. The σ_x^p is a scale parameter. A value of 22 is used for the scale parameter for all the case studies. The

σ_w^2 is the variance of the noise distribution. σ_w^2 may be unknown for the measurement. In that case we should solve a co-optimization problem to estimate $\hat{\sigma}_w^2$ and $\hat{f}$[57], [117]. This is important for our study, because in thermal imaging the noise is usually unknown and an accurate estimate of variance may not be available. To that end an additional ML estimate should be solved in each iteration to estimate $\hat{\sigma}_w^2$ and then use that in equation 5.5 to minimize the main cost function. Therefor $\hat{\sigma}_w^2$ has to be update in each iteration by equation 5.6:

$$\hat{\sigma}_w^p = \frac{1}{N}\sum_{\{i,j\}\in S}(f_i - f_j)^p \tag{5.6}$$

where S is the total number of pixels in the image. After posing the optimization problem, we used Iterative Coordinate Descent (ICD) method to minimize the cost function. The details are available in [117].

In order to test this reconstruction technique, we first start with a numerical "experiment" in which we use ANSYS to create the temperature profile of a 200nm heater line. The dimension and material properties are set according to real device and measured properties as shown in sections 5.2 and 5.3. Next, we filtered the image using the diffraction function of a 100x objective lens with N.A. = 0.75. Finally, we add random Gaussian noise with variance $\hat{\sigma}_w^2 = 20$ (i.e. the signal-to-noise value of ~1.5dB). The blurry noisy image (*numerically produced sub-diffraction temperature map*) are shown in Figures 5.8a. We used the algorithm explained above to reconstruct the thermal image from the sub-diffraction thermal image. This is shown in Figure 5.8b along with the original thermal map. The thermal maps agree very well. A comparison between the cross sections are

shown in Figure 5.8c. The algorithm works markedly in reconstructing the true thermal image.

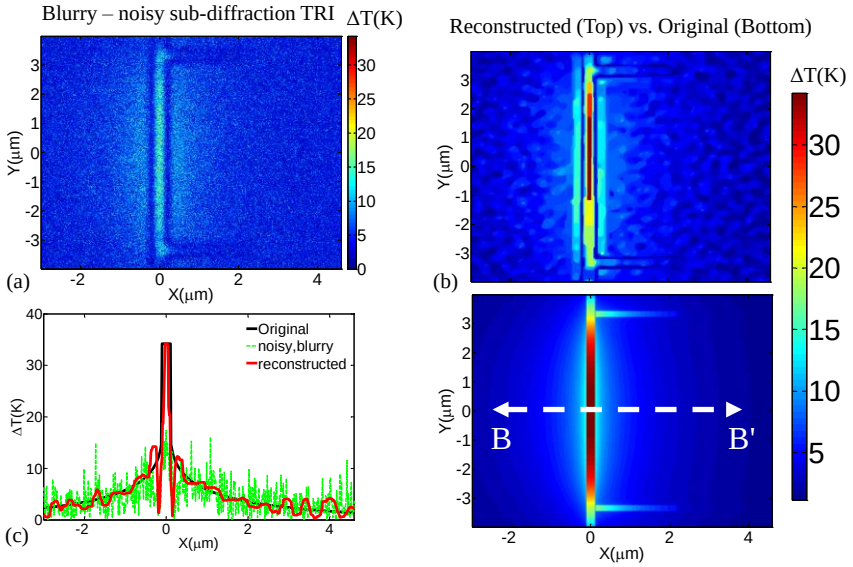


Figure 5.8. Numerical "experiment" to verify the reconstruction algorithm. (a) A noisy-blurry image created by blurring the FEM temperature profile with the diffraction function of 100x objective lens and adding random noise to the image. (b) Reconstructed thermal image (Top) using the MAP estimation framework described in the text along with the original temperature profile (bottom) obtained from FEM. (c) Comparison of cross sections along B-B' shows excellent agreement.

We tested the same algorithm on a set of real thermal images. An example is shown in Figure 5.9. We first measured the thermal image of a 200nm wide heater line (8μm long) with a 100x objective lens. Current applied is 3.65mA. Figure 5.9a shows the experimental thermal image. As it was already mentioned, the resolution limit is 353nm for 100x lens using 530nm LED light. Thus the 200nm heater line is below diffraction. Next, we normalize the thermal image with the measured gold thermoreflectance everywhere (~-2.2×10^{-4}). Performing the reconstruction, we restored the thermal image as it is shown in Figure 5.9b. The reconstructed profile is compared with FEM results and shows excellent agreement. We should emphasize that the true maximum temperature is almost four times

higher than that in the apparent thermal image. This was also measured independently using electrical characterization, confirming the FEM results (See chapter 6 for more details about electrical measurements and the details of the FEM simulations).

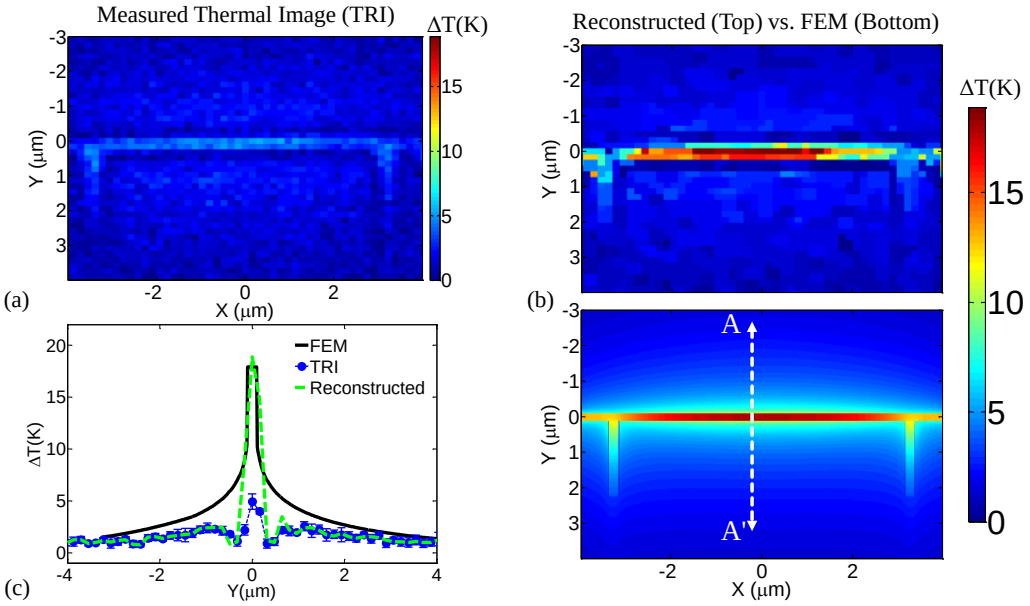


Figure 5.9. Image reconstruction using measured TR images. (a) TR image of the 200nm heater line. The image is calibrated with nominal C_{TR} of gold (~-2.2×10^{-4}). (b) Reconstructed image (Top) using the MAP estimation framework described in the text along with the FEM temperature profile (bottom). (c) Comparison of cross sections along A-A'. Modeling results is verified using independent electrical measurement as it is described in chapter 6.

5.5.Summary

A series of thermoreflectance (TR) thermal imaging measurements on heater lines ranging from 100nm-10μm are performed to systematically quantify the effect of optical diffraction on thermal profile of device with sub-diffraction features. The results suggest that TR imaging microscopy is able to measure the temperature of devices with features smaller than diffraction limit, although the optical diffraction causes the temperature values be

significantly underestimated. We believe similar to phase shift mask in photolithography [118], having coefficients of thermoreflectance of opposite signs between metal and semiconductor enables observation of metal thermal signal while the standard reflected (DC) image is obscured. In the case of 200nm heater line, a factor of 5 reduction in apparent maximum temperature is observed. A combined numerical-analytical modeling procedure is developed to solve the forward problem and successfully reproduce the experimental temperature rise in nanoheater lines. This technique is based on the Gaussian approximation of the diffraction function which is used as a blurring function and convolved with numerical FEM results. The good fit with experimental results suggests that, this blurring function can be utilized in a deblurring (reconstruction) algorithm to extract an accurate temperature profile of devices with sub-diffraction features from thermoreflectance thermal images. To that end a maximum-a-posteriori (MAP) estimation framework with Markov Random Field priors (with p=1) is developed to reconstruct the true temperature map from apparent measured temperature map. Case studies including a *numerical simulation* and a real thermal imaging measurement results are tested. Excellent agreement between reconstructed images and true temperature maps are demonstrated.

In summary, experimental, numerical and analytical modeling were carried out on sub diffraction limit metal lines with feature sizes less than the illumination wavelength as well as the pixel resolution of the thermoreflectance imaging camera. All the experimental case studies shown in this work can be independently characterized and calibrated using electrical measurements, and thus they do not need to solve the inverse problem to obtain the true temperature. However, in reality and for actual temperature of small scale

electronic devices such as submicron gates in transistors, calibration and independent measurement may not be applicable, and this technique is a powerful tool for accurate non-contact temperature measurement of nanoscale features.

6. STUDY OF SUBMICRON HEAT TRANSPORT IN INGAAS

6.1.Introduction

Study of thermal transport is a crucial part of electronic device design and optimization. The governing law of heat conduction, Fourier equation, has been the subject of scrutiny since the size of electronic, thermoelectric, optoelectronic devices has reached submicron scales. It has been predicted and shown that as the thermal transport length scales are reduced and they become on the order of the phonon mean free paths, Fourier diffusion equation fails to explain the thermal behavior [35]–[38]. It was posited that nonlocal and nonequilibrium effects due to phonons carrying heat ballistically away from the heat source are the main cause of the departure from the Fourier law. Experimental investigations using laser based Time Domain Thermoreflectance (TDTR) [43]–[47] and Frequency Domain Thermoreflectance (FDTR) [48] techniques, using Coherent Soft X-ray beams [49], as well as using 2ω/3ω electrical measurement [50], have probed the non-diffusive thermal transport in semiconductor materials and alloys.

In these measurements, the thermal transport length scale (characteristic or important length scale) was varied directly by changing the width of the heat source on the substrate under study [44], [47], [49], by varying the frequency of the thermal excitation and in turn thermal penetration depth in the material under investigation [43], [48], by varying both the width and frequency of the heat source [119], and by altering the distance between

heater and thermometer in a 2ω/3ω electrical measurement [50]. Evidences of non-diffusive thermal transport manifested themselves by an apparent reduction in the thermal conductivity of the material under study. This reduction in thermal conductivity implies a hotter device at the same power density at submicron scale which in turn suggests that thermal management and better understanding of the underlying physics is crucial. Nevertheless, most of these measurements are interpreted using a modified Fourier theory that is Fourier heat equation with an *effective* thermal conductivity. Then modified thermal conductivity is used to reconstruct the phonon mean free path (mfp) distribution[120]. In this picture, the apparent reduction in the measured thermal conductivity is attributed to the fact that phonons with mfp longer than the characteristic length do not contribute to the heat transport.

Maassen and Lundstrom solved the Boltzmann Transport Equation (BTE), and showed that the Fourier equation can capture the physics of the problem given the appropriate boundary condition imposed. They argue that the apparent reduction in thermal conductivity is due to reduction in temperature gradient and not the actual thermal conductivity [121], [122]. This analysis is based on a generalization of the gray model for phonon transport in which different phonon modes propagate in the material independently.

Vermeersch, et al. demonstrate that the quasiballistic transport in semiconductor alloys is governed by Lévy superdiffusion. In this picture the thermal transport inside the alloys is governed by Lévy dynamics and not the random Brownian motion. It is shown that the apparent reduction observed in thermal conductivity is a direct trace of superdiffusive Lévy transport [46], [58].

Using a beam-offset TDTR measurement[119], Wilson and Cahill showed that the failure of Fourier theory in silicon is anisotropic, that is the effective thermal conductivity extracted in in-plane and cross-plane directions are different. They were able to use an anisotropic Fourier model with modified conductivities to fit the experimentally measured data. They also developed a two channel model in which they separate diffusive and ballistic phonons to explain non-diffusive thermal transport and the departure from Fourier theory.

Here we present for the first time a detailed experimental investigation of nanoscale thermal transport in InGaAs using full field thermoreflectance imaging both in steady-state and in transient. This study suggests that modified Fourier theory ceases to explain the full thermal distribution of nano-scale size heat sources. III-V semiconductors and in particular InGaAs are of interest for high electron mobility transistors (HEMT) and continuation of Moore's law to sub-nanometer devices [123]. They are also a suitable candidate for non-equilibrium and nonlinear thermoelectric refrigerators [34], [124]. We employed transient and static thermoreflectance (TR) thermal imaging technique to capture full 2D temperature map of the InGaAs with heater line sources ranging in width from 100 nm to 10μm. Transient results included fast response with 50ns time resolution. Compared to TDTR, FDTR, X-ray beam and 3ω, TR imaging provide two advantages: 1) It provides full 2D temperature map of the device under test, which means we do not rely on a single point average temperature measurement; and 2) The heater line (heat source) is electrically excited, thus the input power could be accurately measured and the temperature of the metal line can also be extracted both using thermoreflectance (using the reflected LED

light from the sample under study) as well as via 4-probe electrical resistivity measurement. Thermoreflectance Thermal Imaging (TRI) offer a more direct probing of the semiconductor thermal response, and bypasses the complexities of the optical techniques such as photon-electron-phonon processes in Al transducer, and complicated time signature: 200 fs laser pulses at 76 MHz repetition at 1-10 MHz modulation, which could potentially distort interpretation of non-diffusive effects.

Four main results are observed. First, in the widest heater line, that is 10μm, the full 3D finite element model based on Fourier diffusion equation can accurately predicts the entire temperature distribution on top of the heater line as well as heat spreading and 2D temperature distribution on the substrate. The experiments and theory match both in static and transient tests. Second, as the width of the heater line is decreased, the maximum temperature measured in the experiment exceeds that of the Fourier prediction. This corresponds to a *lower effective thermal conductivity* of InGaAs film in the Fourier model and this is consistent with the previous literature. Although the extracted *effective* thermal conductivity can be used in the Fourier model to fit the temperature distribution and average temperature measured on top of the heater line, the full temperature distribution on the substrate cannot be explained. We observe that as the size of heat source decreases, Fourier model overestimate significantly the temperature distribution on the substrate. This suggests that a *larger effective thermal conductivity* in the Fourier model is needed to fit the substrate temperature distribution. The two opposite trend for thermal conductivity implies the shortcomings of the modified Fourier model to fully explain the non-diffusive thermal transport observed in InGaAs. Fourth, transient TR imaging results shows that the

thermal response to a 1μs electrical heating pulse for smaller devices leads to larger temperatures than that predicted by Fourier, but the temperature decays faster than Fourier after the excitation pulse ends. The former observation could be explained using a smaller effective thermal conductivity for InGaAs while the latter suggests a larger thermal conductivity. So practically, a modified Fourier model with a single effective thermal conductivity does not provide a full explanation for all these observations. It is worth mentioning that Wilson and Cahill [119] argue that the observation of thermal conductivity reduction as a function of frequency [43] occurred merely in the cross-plane direction because in-plane heat current was negligible. Vermeersch et al. [46] have suggested that microscale thermal transport in InGaAs films studied here, is governed by fractal Lévy superdiffusion. A characteristic feature of the associated single pulse temperature response is that they were higher than conventional diffusive counterparts in the immediate vicinity of the heat source, but lower a short distance away. The experimental observations presented here are thus in qualitative agreement with those theoretical findings**.** Quantitative Lévy-type modeling of our measurement configuration, with detailed account for the spatial and temporal signatures of the heat source, will be a topic for future investigation.

The rest of this chapter is organized as follows. In section 6.2, we explain briefly our device fabrication and experimental setup. The FEM modeling is described in section 6.3.

The measurement results are presented and analyzed in section 6.4. We conclude the chapter in section 6.5.

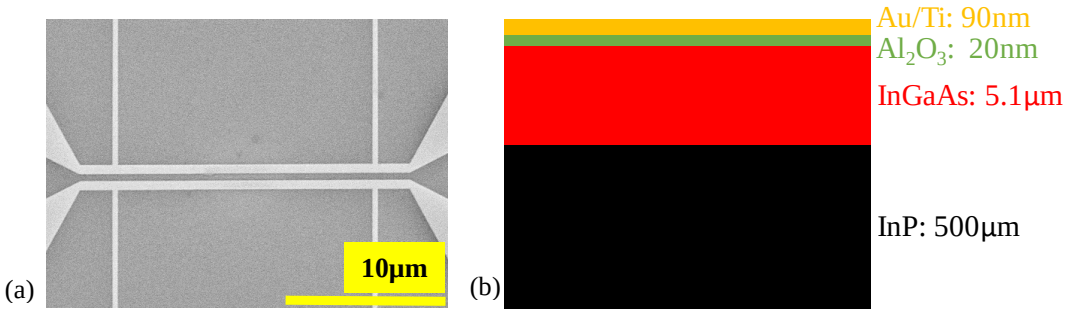


Figure 6.1. Nanoheater lines device structure. a. Scanning Electron Micrograph (SEM) of two 500nm heater lines that are 20µm long each. The gap between the two lines is 500nm. b. Cross section of the structure. Different layers' compositions and dimensions are shown.

6.2.Fabrication and Experimental Setup

A set of heater lines with different widths ranging from 100 nm to 10 µm were fabricated. The details of fabrication are explained in section 5.2. The measured width of the metal thickness was 90 nm. Aspect ratio of each device, i.e. ratio of length to width, was fixed to 40. Four large contact pads, each 80×80µm^2, were fabricated for each heater line, so that the samples can be probed easily and also the thermal measurement can be further confirmed using electrical measurement of the heater resistance. Two similar heater lines were placed in parallel next to each other with distances (gap sizes) of 300 nm, 500 nm and 20 µm. In this case one of the heater line works as a heater and the other heater line serve as thermometer both in electrical and thermoreflectance thermal imaging measurements. Scanning Electron Micrograph (SEM) of two 500 nm lines that are 20µm long and spaced

500 nm, along with the vertical cross-section of the structure is shown in Figures 6.1a and 6.1b. The actual dimensions of the devices might slightly be different from the nominal values due to small variations in electron beam writing. In the analysis we use the actual heater dimensions directly obtained from SEM. Temperature dependent four-probe resistance measurement (IVT) is used to extract the electrical resistivity and thermal conductivity (via Wiedemann-Franz law), as well the temperature dependent resistivity (TCR) of the gold metal lines. From the resistivity (ρ) measurement at different electrical currents and by knowing the TCR (α), we obtain the average temperature rise in the metal line (ΔT) due to increase in electrical current ($\rho=\rho_0(1+\alpha\Delta T)$) [38]. ρ_0 is the nominal resistance of the device and ρ is the resistance at larger currents.

The thermoreflectance (TR) thermal imaging microscopy is used the measure the full 2D spatial temperature profiles of heater lines of different sizes both in steady-state and transient regimes [27], [29], [102]. Time Domain Thermoreflectance (TDTR) [125] and 3ω electrical characterization [126], [127] techniques are used on the same sample to extract the thermal conductivities of thin film InGaAs, metal/semiconductor thermal interface resistance and the substrate thermal conductivity. TDTR is performed on the gold contact pad with an area of 80×80μm^2, and 3ω is performed on the 10×300 μm^2 device which is the largest heater line. Thermal and electrical conductivities of gold as well as its TCR are obtained from current-voltage measurements at different ambient temperatures (IVT). The effective thermal conductivity of the oxide insulation later is measured to be 0.65 ± 0.07 W/m-K. The thermal conductivity of InGaAs and InP were also measured to be 5.5 ± 0.4, and ~70 W/m-K.

6.3.Finite Element Modeling

ANSYS APDL Finite Element Modeling was used to perform full 3D steady-state and transient modeling of the devices. Material properties for different layers are set according to measured values provided in section 6.2. Thermal and electrical conductivity of gold as well as its TCR are obtained from IVT measurements. All material properties are provided in Appendices D and E. Heat capacity and mass density for transient modeling are also set according to literature values for Au [128], Al_2O_3 [129], InGaAs and InP [130]. Over a million elements are used in the full 3D FEM model to ensure accuracy of the modeling. A semi-analytical model is developed based on the equivalent thermal resistance network and the full 3D FEM modeling. The semi-analytical model depends only on thermal conductivities of each layer and serves as an optimization tool to find the InGaAs thermal conductivity that best fit experimental results as we decrease the size of heater lines. All other material properties are assumed to be independent of the device width. The extracted thermal conductivity is then input to the full 3D model to obtain full temperature distribution of heater lines and compared them with experimental results obtained from TR Imaging.

6.4.Results and Discussions

Figure 6.2a shows the temperature profiles obtained by thermoreflectance thermal imaging for a 10μm wide heater line that is in parallel with another 10μm wide heater line with a gap size of 3μm. Temperature maps at different current levels were measured. In Figure 6.2b, the average temperature rise at the top of the heater line, that is measured using TRI, is compared with temperature change measured using IVT and also with the FEM (Fourier

model) results obtained from ANSYS. Measured temperature cross section along the heater line is also shown in Figure 6.2c and compared with the FEM model. It is apparent from these figures that the temperature distribution calculated by FEM at the top of the heater lines follows the experimental results. Additionally, transient measurements were performed using electrical pulses of different widths and duty cycles. A typical transient response was obtained using 1μs electrical heating pulse with 5% duty cycle. The temperature map of the heater line was recorded every 100ns. The average temperature evolution over time is plotted in Figure 6.2d and is compared with FEM simulations. Results shown in Figure 6.2b-d suggest that the experimental data and the Fourier FEM model are in very good agreement. In these figures, FEM was performed using the material and device parameters provided in section 6.2. A thermal conductivity of 5.35W/m-K for InGaAs was used.

IVT and TRI measurements were performed on devices ranging in widths from 10μm down to 100 nm. The same FEM model was also used to calculate the temperature profile of these devices. Parameters of the model are set based on the previous section except for each metal line, properties such as electrical and thermal conductivities, as well as TCR are individually calibrated using the IVT measurements. All the material properties are listed in supplementary materials. As the width of the devices decrease, the Fourier based FEM model fails to accurately predict the temperature profiles of the heater lines and underestimates the temperature change (ΔT_{FEM}) compared to those measured from experiments (ΔT_{Exp}). This could be explained by a change in “effective” thermal conductivity of the InGaAs thin film consistent for both static and transient regimes. Figure

6.3a-d, shows the experimental and modeling results for a 400 nm wide heater line. Static Temperature profile is shown in Figure 6.3a. Temperature rise is plotted against current in Figure 6.3b for IVT, TRI and FEM. It is evident from this figure that using the nominal thermal conductivity of InGaAs (~5.4W/m-K) results in about 18% underestimation compared to the experimental data. Instead, a low effective thermal conductivity of ~4.2W/m-K for InGaAs must be used in the model to match the top surface temperature.

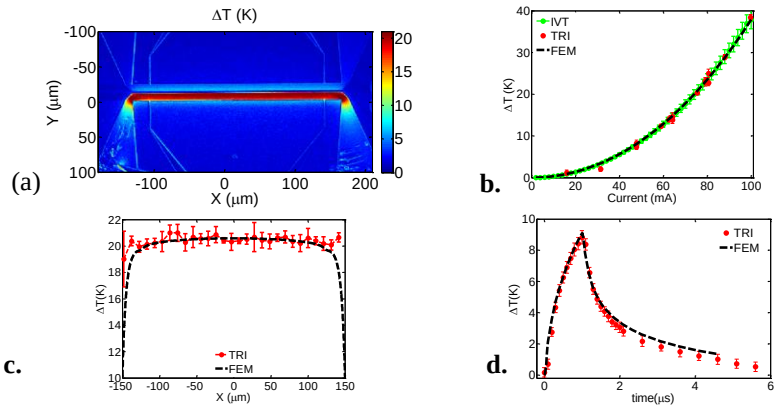


Figure 6.2. Thermoreflectance thermal imaging (TRI) and IVT results for 10μm wide heater line. **a.** Temperature map (thermal image) of the devices is shown. Electrical current is 74.5mA. **b.** Average temperature change at different electrical currents on top of the heater line extracted using IVT, TRI and Finite Element Modeling (FEM). **c.** Comparison of temperature cross sections along the heater line obtained in the experiment (TRI) and FEM. **d.** Comparison of maximum transient temperature on top of the heater line using TRI and FEM. A 1μs current pulse with 5% duty cycle was applied to the heater and the temperature profile was recorded every 100ns. The average temperature near the center of the line is compared to that of obtained from FEM. Good agreement between the model and the experiment in all of the results is evident. Thermal conductivities of InP, InGaAs, and the oxide film are assumed to be 70W/m-K, 5.35 W/m-K and 0.65 W/m-K, respectively.

Figure 6.3c shows the cross section along the top of the 400 nm wide heater line for TRI and the comparison with Fourier models with both nominal (~5.4 W/m-K) and adjusted (~4.2 W/m-K) thermal conductivities for InGaAs. The evolution of transient temperature

change is plotted in Figure 6.3d. These results suggest that in the case of Fourier model with modified effective thermal conductivity, both shape and magnitude of the distribution of the temperature change *at the top of the heater lines* follows the experimental results.

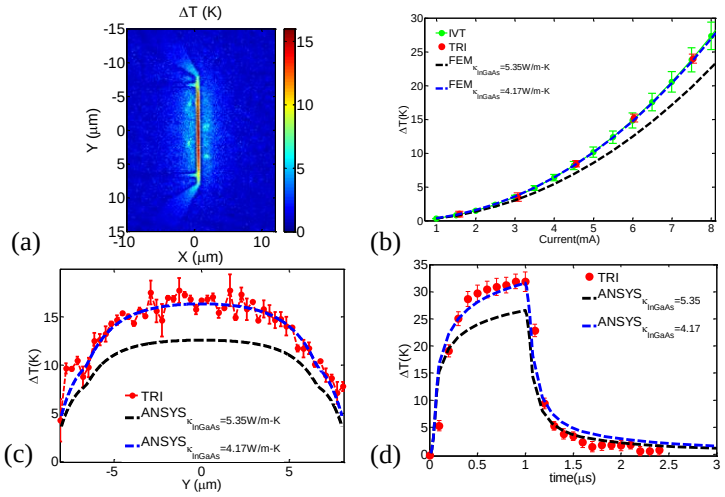


Figure 6.3. Thermoreflectance thermal imaging (TRI) and IVT results for 400nm heater line. **a.** Temperature profile of the 400nm line under 6mA electrical current excitation. **b.** Measured average temperature change at different currents using IVT, and TRI. FEM, based on Fourier theory, with InGaAs thermal conductivity of 5.35 W/m-K underestimates the temperature by ~20% (black curve). By reducing the thermal conductivity of InGaAs to 4.17W/m-K, the FEM result on top of the heater follows the experimental data within 0.5% (Blue curve). **c.** Temperature cross section along the top of the 400 nm heater line. It is evident from this figure that by modifying the thermal conductivity, theory and experiment can match. d**.** Transient temperature at the top surface as a function of time (3 μs duration with 100 ns steps) is plotted for TRI and FEM model with different thermal conductivities. A 1μs electrical pulse with 5% duty cycle (9.16mA) was applied to the heater line.

A similar approach for all the device sizes (0.1-10 μm) was taken. The deviation from Fourier model (in %) and the corresponding modified effective thermal conductivity, that compensates the deviations, are plotted against width of the device in Figure 6.4a and 6.4b. At 200nm and 100nm widths, InGaAs thermal conductivities of 4.2 W/m-K and 2.7 W/m-K give the best fit which is about 21% and 50% reduction compared to the bulk value.

These results suggest that as the physical dimension of heat source is reaching phonon mean free path in InGaAs, we may use a reduced apparent thermal conductivity to fit the experimental results on top of the heater line [43], [44], [131], [132].

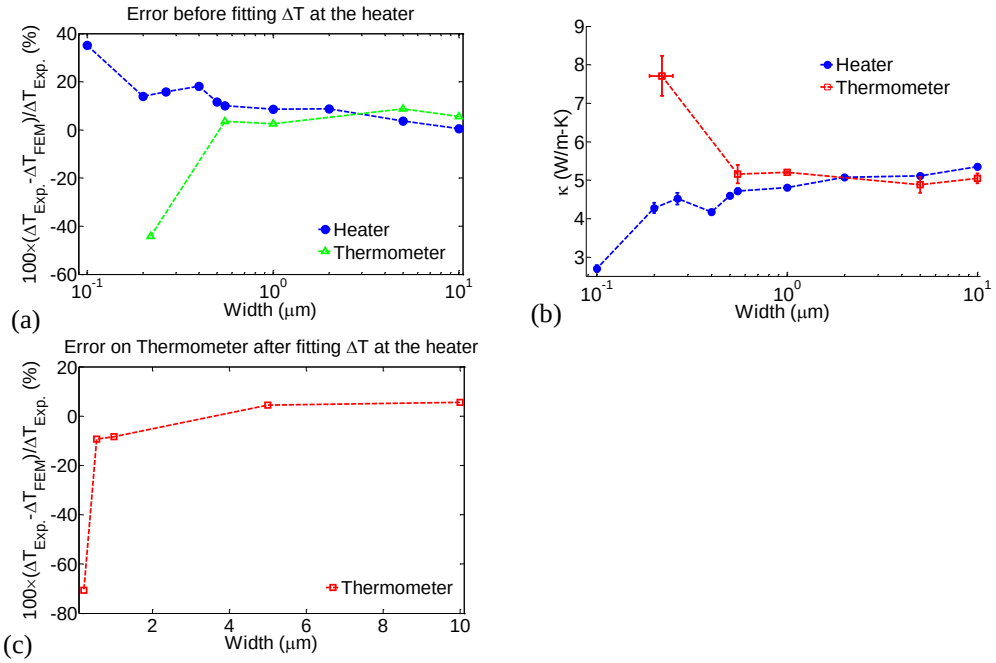


Figure 6.4. Differences between temperature changes obtained from the Fourier diffusive model and the experimental temperatures and corresponding Apparent Thermal Conductivities to fit the temperature at the heater and the thermometer. **a.** The difference in % between the measured temperature on top of the heater line and that predicted by Fourier diffusive model (FEM) using nominal InGaAs thermal conductivity (κ=5.4W/m-K, blue circles). The same comparison between theory and experiment for the temperature data on top of the thermometer line (inactive heater used as a sensor) (green triangles; using an isotropic FEM with κ=5.4W/m-K). **b.** Apparent thermal conductivity needed to fit the temperature of the heater (blue circles) and the thermometer (red squares). It is important to emphasize that, similar to Wilson and Cahill [119], we were able to fit the full temperature distribution with an isotropic Fourier model. The numbers in Figure b, only independently matches the average temperature at the top and on the sensor. The entire distribution can be fitted with an anisotropic thermal conductivity for InGaAs if cross-plane value of 10 W/m-K and in-plane thermal conductivity of ~1.5W/m-K is assumed which are quite surprising numbers. **c.** Comparison between theory and experiment on top of the thermometer (red squares), but this time a modified κ (according to the blue dots in part b) are assumed to fit the ΔT on the heater.

Despite being able to predict the temperature change and distribution at the top of the heater lines by modifying the thermal conductivity of InGaAs alloy film in the Fourier model, the temperature distribution on the substrate, tail of distribution next to heater line and the temperature distribution on a thermometer line within few microns of the source, cannot be predicted by the same model. This is shown in Figure 6.5. Figures 6.5a1-a5 show the temperature profiles of heater lines fabricated within few microns of a thermometer line. Electrical current was sent through the heater line to change its temperature and the neighboring line serves as a thermometer. The dimensions of the heater lines as well as the thermometer lines are given on top of each device temperature map. Color scale is adjusted in Figures 6.5a1-a5 to clearly show heat spreading in the substrate and through the thermometer lines. The temperature cross sections of different devices along A-A', perpendicular to the lines in direction of y-axis, are plotted in Figures 6.5b1-b5. A ~5.4 W/m-K for 10μm device, and a 4.5 W/m-K for the 265nm device were used in the FEM model, which corresponds to the same number shown in Figure 4b. It is evident that as the width of the heater lines decreases, the deviation of the temperature distribution on the thermometer lines next to the device from the modified Fourier model increases. In fact, this effect is magnified in Figures 6.5c1-c5. At 265nm, the average temperature on the thermometer line is ~80% lower than that of predicted by modified Fourier model. Hence, Figure 6.5 illustrates that although by using an *apparent* thermal conductivity in the Fourier model we are able to predict and fit the experimental results, the full distribution of temperature profile cannot be predicted. Figure 6.4a and c, plot the deviation from Fourier model to explain the tails of the temperature distribution, both before fitting the

temperature at the top surface (green curve in Figure 6.4a), assuming a constant thermal conductivity of ~5.4W/m-K in the model, and after fitting the temperature distribution at the top (Figure 6.4c), using *apparent* cross-plane thermal conductivities presented in Figure 6.4b (blue circles) for the model. Moreover, the ratio of the average temperature change on the thermometer line between modeling and experiment was used to calculate the effective in-plane InGaAs thermal conductivity that fits experimental results. This effective thermal conductivity is plotted as a function of width on the left axis in Figure 6.4b (red squares). However, using these effective thermal conductivities for InGaAs in the Fourier model results in as large as %35 difference between experimental and modeling results at the top. The anisotropic behavior is also apparent in temperature dependent measurements. Figures 6.6a and b summarize the results measured using TR for 10μm and 1μm heater lines. In the case of 10μm heater line, the anisotropy is weak and becomes more evident as we reach below 100K. For 1μm heater line, there is an increase in anisotropy as temperature decreases up to 130K. It is interesting to note that below 150K, the cross-plane thermal conductivity of InGaAs for 10μm line drops faster than the 1μm line and the thermal conductivities obtained for the 1μm line are larger. This contradicts the room temperature observations in which as the size of heat source decreases the apparent thermal conductivity decreases. This requires further analysis as Al_2O_3 thermal conductivity may reduce sharply at low temperatures which could complicate the extraction of the interface thermal

resistance. At smaller scales, as we show in the appendix C, the interface becomes more important and its temperature dependence should be extracted carefully.

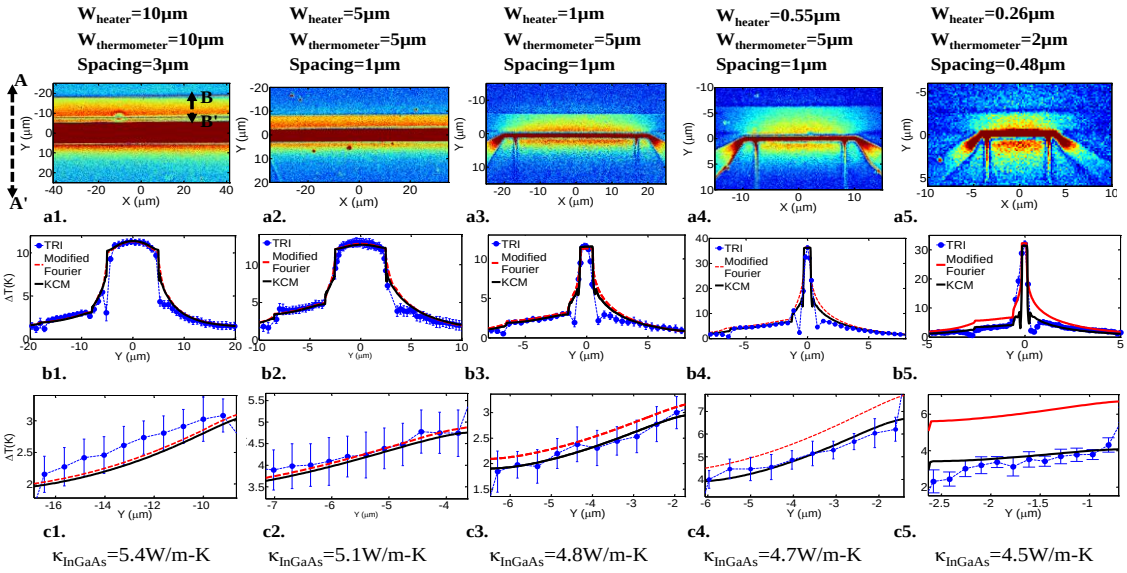


Figure 6.5. Experimental 2D temperature profiles on top of the heater line, thermometer line and the substrate. **a1-5.** Temperature profiles of two heater lines in parallel. One acts as the heater and the other act as a thermometer (known thermoreflectance coefficient and temperature coefficient of resistance for gold). The dimensions are shown on top of each image. The aspect ratio (length/width) for all heaters is 40. Color scales are adjusted so that the heating can be seen more clearly. **b1-5.** Temperature profile across the y-axis (perpendicular to the heater line). In the modified Fourier model, the Thermal conductivity of substrate is adjusted so that the numerical maximum temperature at the top agrees with experimental results. The black curve shows the preliminary non-local kinetic-collective simulation which takes into account the superdiffusion (this was developed by our collaborator at Univ. of Barcelona and it should be finalized shortly). **c1-5.** As the size of heater line decreases, the temperature obtained from modified Fourier model for the thermometer line exceed the measurement at that location. In particular, for the 265nm heater line, the average temperature on the thermometer obtained from Fourier model exceeds the measured average temperature by ~80%. NL-KCM, on the other hand, agrees with the experimental results both on top of the heater as well as for the tails of the temperature distribution.

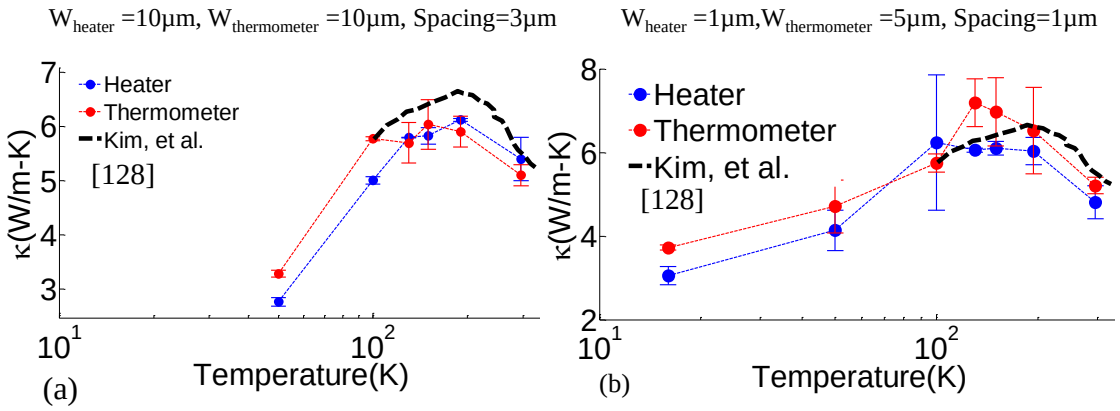


Figure 6.6. Temperature dependent thermal conductivity of InGaAs. (a) For 10µm heater line. (b) For 1 µm heater line. The blue dots are extracted based on the temperature at the top of the heater line (cross-plane κ), while the red dots are extracted based on measurement on the neighboring thermometer line (in-plane κ). The black dashed line is from reference [133].

Not only the modified Fourier theory cannot explain the full steady-state temperature distribution, but also the transient response of the heater lines shows discrepancies between the experiment and modified Fourier model. This is shown in Figure 6.7. The temperature evolution over time in response of a 1µs electrical pulse with 10% duty cycle is plotted in Figures 6.7a-d for 500nm, 400nm, 265nm, and 200nm heater lines, respectively. Interestingly, the rising time of the experimental results with respect to modeling is slower, while the falling time of temperature evolution is much faster. The effect is more prominent as the size of devices decreases. A larger temperature at the top suggest a reduced apparent thermal conductivity, while faster decay suggests the opposite. The inability of the modified Fourier theory to explain the temperature distribution of the heater lines and also their transient temperature evolution, implies that an alternative theory beyond modified Fourier theory is needed to fully describe the experimental results.

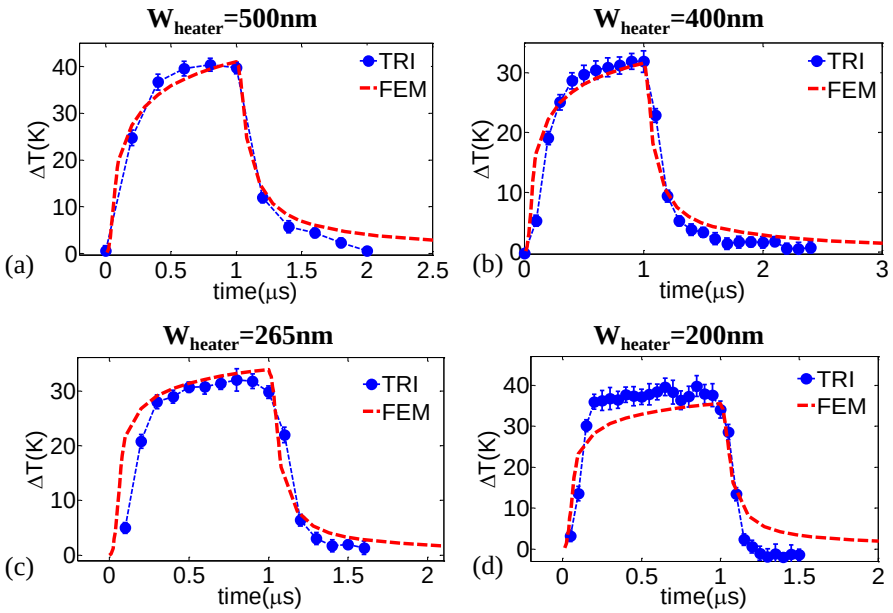


Figure 6.7. Deviations between Fourier diffusive model and the experimental transient temperature response. For smaller heater lines, the experimental temperature rises slower and decays faster compared to the Fourier theory (whose thermal conductivity is adjusted to get accurate temperature on top of the heater). Temperature evolution against time is shown for **a.** 500nm; **b.** 400nm; **c.** 265nm; and **d.** 200nm wide heater lines.

6.5. Conclusions

We utilized thermoreflectance thermal imaging (TRI) with high spatial and temporal resolutions as well as IVT electrical characterization technique to study heat transport in InGaAs thin films as a function of the size of the localized heat source. It was shown that the temperature change on top of the heat source increases beyond that predicted by Fourier model for small scale heaters. 20-30% increase in temperature beyond Fourier prediction was observed for heater lines with width below 400 nm. This effect can be addressed by modifying the InGaAs thermal conductivity and using an *apparent* thermal conductivity in Fourier model. However, it is demonstrated that modifying the thermal conductivity of

InGaAs could only fit the temperature at the top of the heater line and not the full temperature distribution around the heater. About 80% difference between modeling (modified Fourier) and experiment was observed on top of 2μm thermometer line placed at a distance of 500nm from 265nm heater line. The trend is different below 150K and this needs further investigation. Additionally, transient temperature response of the heater line under short electrical pulse excitation exhibits deviations from modified Fourier model both in rising and falling times. An alternative approach beyond modified Fourier theory that can explain these effects is indispensable.

7. FUTURE WORK

7.1.Effect of Surface Plasmonic Resonances On TR Imaging of Nanoscale devices

While studying thermal transport in nanoscale size devices, as shown chapters 5 and 6 of this dissertation, several optical effects were observed that led us to postulate the possibility of surface plasmonic effects on the thermal images. These results are summarized in Figures 7.1 and 7.2. In Figure 7.1a and b the SEM and the thermal images of two 160nm heater lines that are spaced 284nm apart from each other are shown respectively. Similar to the approach presented in chapter 5 we blurred the FEM results with the appropriate diffraction function in order to compare the measured temperature profile with that obtained using finite element simulation. Two main points are observed: 1) The simulated maximum apparent temperature is much lower compared to the experimental results; And 2) The location of the peak in experiment happens somewhere in the gap and not on top of the active metal line. Both effects can be explained if a much larger C_{TR} (a factor of 8 larger) in the gap is used compared to the C_{TR} on the rest of the substrate. This increase in the C_{TR} in the gap could be due to plasmonic resonances at the wavelength we are using for the measurement. A full systematic study, TR imaging with different wavelengths and plasmonic electromagnetic simulations for different device dimensions are needed to fully understand the physics.

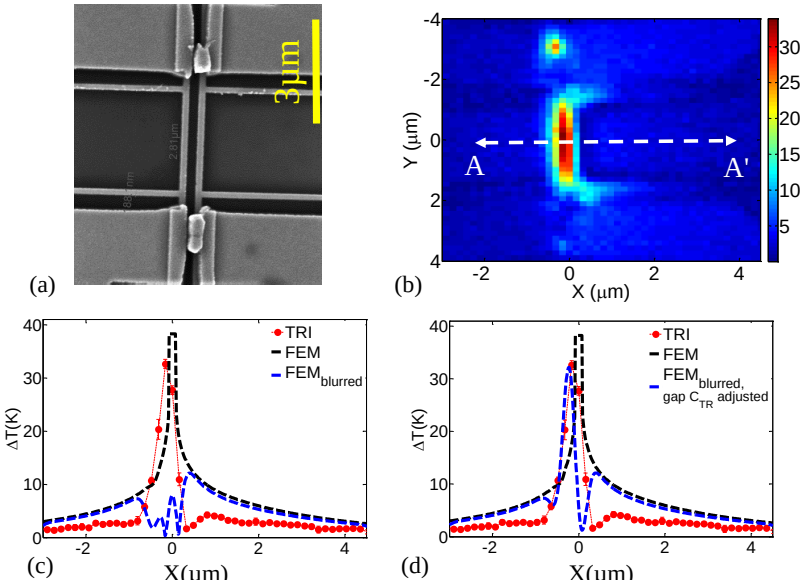


Figure 7.1. Thermal imaging of two 160nm heater lines that are spaced 284nm apart from each other. (a) SEM image of the two heater lines. (b). Thermal image obtained. Current sent through the right heater line. (c) comparison of cross section along A-A'. Red curve is the experimental data, black curve is the FEM results, and blue line is the blurred results. (d) Blurred results with modified CTR in the gap to match the blurred results to experimental measured thermal image.

The second observation is shown in Figure 7.2. Phase thermal images for 1µm and 200nm wide heater lines at 530nm and 455nm illumination wavelengths are shown in Figure 7.2a-d. One notices that as the size of the device is decreased from from 1µm to 200nm, the sign of C_{TR} for the metal changes at 455nm. This means the metal at 455nm has a positive CTR in 1µm device and a negative CTR at 200nm device. This effect was not observed at 530nm and 660nm (not shown here). This suggest an optical effect occurring at these wavelengths that cannot be neglected and will affect the thermal imaging response. These optical effect could adversely influence the TR imaging results. Therefore, treating plasmonic effects

and incorporation of nano-photonic simulations are important to accurately extract nanoscale temperature profile and quantify non-Fourier heat conduction.

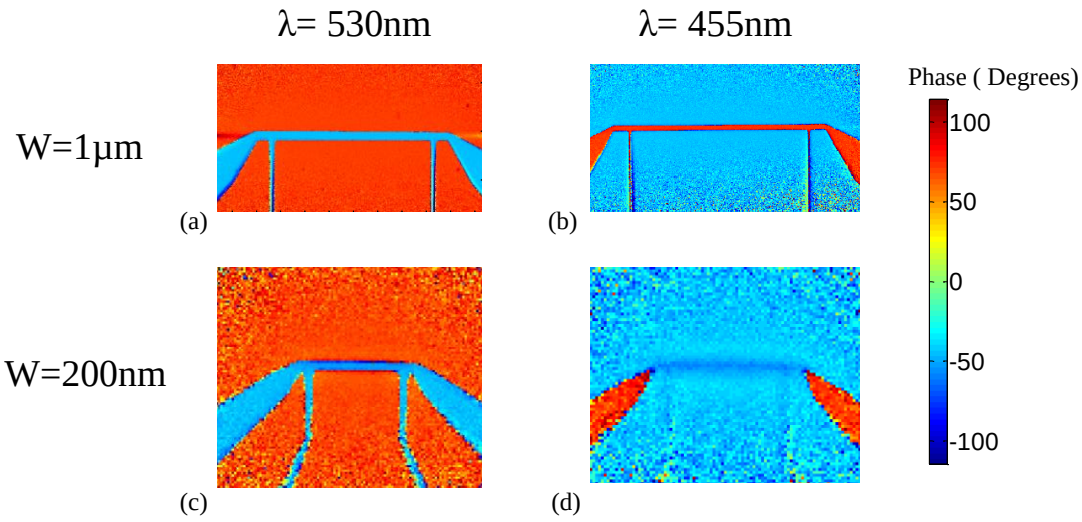


Figure 7.2. Change in sign of CTR due to the size of metal line at certain wavelengths. Phase image at (a) W=1µm, λ=530nm, (b) W=1µm, λ=455nm, (c) W=0.2µm, λ=530nm, and (d) W=0.2µm, λ=455nm. It is clear in image (d) the 200nm heater line and the gold pads have different C_{TR}.

7.2.Sub-diffraction Thermal Imaging and Thermal Image Reconstruction for State-Of-The-Art Nanoscale Electronic Devices

In Chapter 5, we described how image reconstruction techniques could be employed to extract accurate temperature profiles of sub diffraction-limit device sizes. The case studies are based on nanoscale heater lines that could be independently characterize electrically in order to extract their average temperature and thermal properties. In fact, these independent studies helped us to verify the FEM modeling results are true and consequently

reconstruction is done accurately. However, in reality many applications deal with nanoscale devices that cannot be locally characterized or calibrated independently. Thermal imaging is a powerful tool for non-contact thermal measurements in these type of devices. Examples include, study of hot carrier injection as a major thermal reliability concern in short channel MOSFETs [113], self-heating due to ballistic transport in nanoscale transistors [53], thermo-migration in high power transistors and interconnects, reliability of heat-assisted magnetic recording (HAMR) devices. In these examples, very high localized temperatures could result in a change in material properties or even device degradation, and therefore an accurate measurement of temperature is crucial to ensure optimized device performance and operation. Due to sub-diffraction features in these devices, thermal imaging could only be used to study the shape of the hotspot, uniformity and/or non-uniformity of the temperature profile, and rough estimate of temperatures of the devices under study. Advanced image/temperature reconstruction techniques proposed in this dissertation could be extremely useful in thermal imaging of state-of-the-art nanoscale devices. Detailed analysis of the accuracy of the reconstruction technique as a function of material properties, signal to noise ratio or characteristics of the optical imaging system will be required.

7.3.Study of Non-Local and Collective Heat Transport in Single Crystal Materials such as Silicon

In chapter 6, we used TR imaging to study non-diffusive thermal transport in InGaAs alloy. The systematic procedure described in chapter 6 can also be used to study thermal transport in silicon at room and at cryogenic temperatures. It is predicted that at low and mid

temperature range, where normal phonon scattering processes are prominent, phonon transport is in collective regime. The main role of normal scattering is redistributing the conserved momentum among all the phonon modes. Thus, any collision (resistive or not) impacts the collection of phonons i.e. phonons transport behaves as a fluid rather that transport by independent channels. This will affect the measured thermal conductivity. Some initial modeling results shows that as the size of heat source reduces these effects become more prominent and the shape of temperature distribution will be very different from that of obtained from Fourier model. The systematic procedure described in chapter 6 along with thermoreflectance thermal imaging technique provide a very nice platform to study the full distribution of heater sources of different sizes and in turn observe any collective transport in materials such as silicon. Silicon has been crucial in advancing the electronic industry, and understanding the physics of heat transport in silicon can have implications for future advancement of nanoelectronics.

Beyond the three topics mentioned above, integrating power blurring technique in an architectural simulator can be useful in fast and accurate modeling of integrated circuits and in particular at early stage of routing and design optimization. Nonlinear thermoelectric effect can be further enhanced by studying different materials or by optimizing the band structure to minimize the effect of the Schottky contacts and barrier lowering near the emitter. Stress analysis presented in chapter two provides some initial suggestions for designing a robust TE system for high temperature industrial waste heat recovery applications.

LIST OF REFERENCES

LIST OF REFERENCES

[1] G. E. Moore, “Cramming more components onto integrated circuits,” *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, 1998.

[2] K. Banerjee, S. J. Souri, P. Kapur, and K. C. Saraswat, “3-D ICs: A novel chip design for improving deep-submieroraeter interconnect performance and systems-on-chip integration and systems-on-chlp integration,” *Proc. IEEE*, vol. 89, no. 5, pp. 602–632, 2001.

[3] J. Kong, S. W. Chung, and K. Skadron, “Recent thermal management techniques for microprocessors,” *ACM Comput. Surv.*, vol. 44, no. 3, pp. 1–42, 2012.

[4] S. Chaudhury, “A Tutorial and Survey on Thermal-Aware VLSI Design: Tools and Techniques,” *Int. J. Recent Trends …*, vol. 2, no. 8, 2009.

[5] G. J. Snyder and E. S. Toberer, “Complex thermoelectric materials.,” *Nat. Mater.*, vol. 7, no. 2, pp. 105–114, 2008.

[6] C. J. Vineis, A. Shakouri, A. Majumdar, and M. G. Kanatzidis, “Nanostructured thermoelectrics: Big efficiency gains from small features,” *Adv. Mater.*, vol. 22, no. 36, pp. 3970–3980, 2010.

[7] A. Shakouri, “Recent Developments in Semiconductor Thermoelectric Physics and Materials,” *Annual Review of Materials Research*, vol. 41. pp. 399–431, 2011.

[8] R. Venkatasubramanian, E. Siivola, T. Colpitts, and B. O’Quinn, “Thin-film thermoelectric devices with high room-temperature figures of merit.,” *Nature*, vol. 413, no. 6856, pp. 597–602, 2001.

[9] W. K. W. Kim, S. Singer, A. Majumdar, J. Zide, A. Gossard, and A. Shakouri, “Role of nanostructures in reducing thermal conductivity below alloy limit in crystalline solids,” *ICT 2005. 24th Int. Conf. Thermoelectr. 2005.*, 2005.

[10] V. Rawat, Y. K. Koh, D. G. Cahill, and T. D. Sands, “Thermal conductivity of (Zr,W)N/ScN metal/semiconductor multilayers and superlattices,” *J. Appl. Phys.*, vol. 105, no. 2, 2009.

[11] M. Zebarjadi, K. Esfarjani, A. Shakouri, Z. Bian, J.-H. Bahk, G. Zeng, J. Bowers, H. Lu, J. Zide, and A. Gossard, “Effect of Nanoparticles on Electron and Thermoelectric Transport,” *Journal of Electronic Materials*, vol. 38. pp. 954–959, 2009.

[12] W. Kim, J. Zide, A. Gossard, D. Klenov, S. Stemmer, A. Shakouri, and A. Majumdar, “Thermal conductivity reduction and thermoelectric figure of merit increase by embedding nanoparticles in crystalline semiconductors.,” *Phys. Rev. Lett.*, vol. 96, p. 045901, 2006.

[13] J. M. O. Zide, J.-H. Bahk, R. Singh, M. Zebarjadi, G. Zeng, H. Lu, J. P. Feser, D. Xu, S. L. Singer, Z. X. Bian, A. Majumdar, J. E. Bowers, A. Shakouri, and a. C. Gossard, “High efficiency semimetal/semiconductor nanocomposite thermoelectric materials,” *J. Appl. Phys.*, vol. 108, p. 123702, 2010.

[14] J. M. O. Zide, D. Vashaee, Z. X. Bian, G. Zeng, J. E. Bowers, A. C. Gossard, A. Shakouri, and A. C. Gossard, “Demonstration of electron filtering to increase the Seebeck coefficient in In_0.53Ga_0.47As/In_0.53Ga_0.28Al_0.19As superlattices,” *Phys. Rev. B*, vol. 74, p. 205335, 2006.

[15] S. H. Choday, M. S. Lundstrom, and K. Roy, “Prospects of thin-film thermoelectric devices for hot-spot cooling and on-chip energy harvesting,” *IEEE Trans. Components, Packag. Manuf. Technol.*, vol. 3, no. 12, pp. 2059–2067, 2013.

[16] V. Sahu, Y. K. Joshi, and A. G. Fedorov, “Experimental investigation of hotspot removal using superlattice cooler,” *2010 12th IEEE Intersoc. Conf. Therm. Thermomechanical Phenom. Electron. Syst.*, pp. 1–8, Jun. 2010.

[17] K. Yazawa, A. Ziabari, A. Shakouri, V. Sahu, A. G. Fedorov, and Y. Joshi, “Cooling power optimization for hybrid solid-state and liquid cooling in integrated circuit chips with hotspots,” *13th Intersoc. Conf. Therm. Thermomechanical Phenom. Electron. Syst.*, pp. 99–106, May 2012.

[18] V. Sahu, A. G. Fedorov, Y. Joshi, K. Yazawa, A. Ziabari, and A. Shakouri, “Energy efficient liquid-thermoelectric hybrid cooling for hot-spot removal,” in *28th Annual IEEE Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2012, pp. 130–134.

[19] K. Yazawa and A. Shakouri, “Exergy Analysis and Entropy Generation Minimization of Thermoelectric Waste Heat Recovery for Electronics,” *Proc. ASME 2011 Pacific Rim Tech. Conf. Photonic Syst.*, pp. 1–7, 2011.

[20] K. Yazawa and A. Shakouri, “Cost-Efficiency Trade-off and the Design of Thermoelectric Power Generators.,” *Environ. Sci. Technol.*, vol. 45, pp. 7548–7553, 2011.

[21] K. Yazawa, Y. R. Koh, and A. Shakouri, “Optimization of thermoelectric topping combined steam turbine cycles for energy economy,” *Appl. Energy*, vol. 109, pp. 1–9, 2013.

[22] K. Yazawa and A. Shakouri, “Optimization of power and efficiency of thermoelectric devices with asymmetric thermal contacts,” *J. Appl. Phys.*, vol. 111, no. 2, p. 024509, 2012.

[23] Y. Han, I. Koren, and C. Moritz, “Temperature aware floorplanning,” *Proc Second Work. Temp. Comput. Syst.*, vol. 51, no. 7–8, pp. 927–934, 2005.

[24] E. K. Ardestani, A. Ziabari, A. Shakouri, and J. Renau, “Enabling power density and thermal-aware floorplanning,” in *28th Annual IEEE Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2012, pp. 302–307.

[25] J. N. Reddy, *An Introduction to the Finite Element Method*, Third. McGraw-Hil, 2006.

[26] M. Pedram and S. Nazarian, “Thermal Modeling, Analysis, and Management in VLSI Circuits: Principles and Methods,” *Proc. IEEE*, vol. 94, no. 8, 2006.

[27] M. Farzaneh, K. Maize, D. Lüerßen, J. A. Summers, P. M. Mayer, P. E. Raad, K. P. Pipe, A. Shakouri, R. J. Ram, and J. A. Hudgings, “CCD-based thermoreflectance microscopy: principles and applications,” *Journal of Physics D: Applied Physics*, vol. 42. p. 143001, 2009.

[28] A. A. Shakouri, A. Ziabari, D. Kendig, J. Bahk, Y. Xuan, P. D. Ye, S. Clara, W. Lafayette, D. Y. Peide, K. Yazawa, and A. A. Shakouri, “Stable thermoreflectance thermal imaging microscopy with piezoelectric position control,” in *2016 32nd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2016, pp. 128–132.

[29] B. Vermeersch, J.-H. Bahk, J. Christofferson, and A. Shakouri, “Thermoreflectance imaging of sub 100 ns pulsed cooling in high-speed thermoelectric microcoolers,” *J. Appl. Phys.*, vol. 113, no. 10, p. 104502, 2013.

[30] B. Vermeersch, J. Christofferson, K. Maize, a Shakouri, and G. De Mey, “Time and frequency domain CCD-based thermoreflectance techniques for high-resolution transient thermal imaging,” *2010 26th Annu. IEEE Semicond. Therm. Meas. Manag. Symp.*, pp. 228–234, 2010.

[31] K. Maize, J. Christofferson, and A. Shakouri, “Transient Thermal Imaging Using Thermoreflectance,” in *2008 IEEE Twenty-fourth Annual Semiconductor Thermal Measurement and Management Symposium, Semi-Therm.*, 2008, pp. 55–58.

[32] A. S. K. Yazawa, D. Kendig, K. Maize, “Transient thermal characterization of HEMT devices,” in *IEEE MTT-S International Microwave Symposium*, 2014.

[33] J. Christofferson, K. Yazawa, and A. Shakouri, “Picosecond Transient Thermal Imaging Using a CCD Based Thermoreflectance System,” in *Proceedings of the 14th International Heat Transfer Conference IHTC14*, 2010, pp. 93–97.

[34] M. Zebarjadi, K. Esfarjani, and A. Shakouri, “Nonlinear Peltier effect in semiconductors,” *Appl. Phys. Lett.*, vol. 91, no. 12, p. 122104, 2007.

[35] G. Chen, “Ballistic-diffusive heat-conduction equations,” *Phys. Rev. Lett.*, vol. 86, no. 11, pp. 2297–2300, 2001.

[36] G. Chen, “Nonlocal and Nonequilibrium Heat Conduction in the Vicinity of Nanoparticles,” *J. Heat Transfer*, vol. 118, no. 3, pp. 539–545, 1996.

[37] A. A. Joshi and A. Majumdar, “Transient ballistic and diffusive phonon heat transport in thin films,” *J. Appl. Phys.*, vol. 74, no. 1, pp. 31–39, 1993.

[38] P. G. Sverdrup, S. Sinha, M. Asheghi, S. Uma, and K. E. Goodson, “Measurement of ballistic phonon conduction near hotspots in silicon,” *Appl. Phys. Lett.*, vol. 78, no. 21, pp. 3331–3333, 2001.

[39] M. Hartmann, G. Mahler, and O. Hess, “Existence of temperature on the nanoscale,” *Phys. Rev. Lett.*, vol. 93, no. 8, pp. 1–4, 2004.

[40] M. Hartmann, “Minimal length scales for the existence of local temperature,” in *Thermometry at the Nanoscale: Techniques and Selected Applications*, vol. 47, no. 2, F. Palacio and L. D. Carlos, Eds. Royal Society of Chemistry, 2015, pp. 89–102.

[41] J. Vazquez-Casas and D. Jou, “Temperature in non-equilibrium states : a review of,” *Rep. Prog. Phys.*, vol. 66, pp. 1937–2023, 2003.

[42] A. Sellitto, V. A. Cimmelli, and D. Jou, *Mesoscopic Theories of Heat Transport in Nanosystems*, vol. 6. 2016.

[43] Y. K. Koh and D. G. Cahill, “Frequency dependence of the thermal conductivity of semiconductor alloys,” *Phys. Rev. B*, vol. 76, no. 7, p. 075207, Aug. 2007.

[44] A. J. Minnich, J. A. Johnson, A. J. Schmidt, K. Esfarjani, M. S. Dresselhaus, K. A. Nelson, and G. Chen, “Thermal conductivity spectroscopy technique to measure phonon mean free paths,” *Phys. Rev. Lett.*, vol. 107, no. 9, pp. 1–4, 2011.

[45] L. Zeng, K. C. Collins, Y. Hu, M. N. Luckyanova, A. A. Maznev, S. Huberman, V. Chiloyan, J. Zhou, X. Huang, K. A. Nelson, and G. Chen, “Measuring Phonon Mean Free Path Distributions by Probing Quasiballistic Phonon Transport in Grating Nanostructures,” *Sci. Rep.*, vol. 5, no. August, p. 17131, 2015.

[46] B. Vermeersch, A. M. S. Mohammed, G. Pernot, Y. R. Koh, and A. Shakouri, “Superdiffusive heat conduction in semiconductor alloys. II. Truncated Lévy formalism for experimental analysis,” *Phys. Rev. B*, vol. 91, no. 8, pp. 1–7, 2015.

[47] A. A. Maznev and J. A. Johnson, “Onset of nondiffusive phonon transport in transient thermal grating decay,” *Phys. Rev. B - Condens. Matter Mater. Phys.*, vol. 84, no. 19, 2011.

[48] K. T. Regner, D. P. Sellan, Z. Su, C. H. Amon, A. J. H. McGaughey, and J. Malen, “Broadband phonon mean free path contributions to thermal conductivity measured using frequency domain thermoreflectance.,” *Nat. Commun.*, vol. 4, p. 1640, 2013.

[49] M. E. Siemens, Q. Li, R. Yang, K. K. A. Nelson, E. H. Anderson, M. M. Murnane, and H. C. Kapteyn, “Quasi-Ballistic thermal transport from nanoscale interfaces observed using ultrafast coherent soft X-ray beams,” *Nat. Mater. Lett.*, vol. 9, pp. 26–30, 2010.

[50] A. T. Ramu, N. I. Halaszynski, J. D. Peters, C. D. Meinhart, and J. E. Bowers, “An electrical probe of the phonon mean-free path spectrum Thermometer wire Heater wire Material under test Electrical isolation layer,” *arXiv Prepr. arXiv1602.00381*, pp. 1–11, 2016.

[51] T. Favaloro, J. Suh, B. Vermeersch, K. Liu, Y. Gu, L. Chen, K. X. Wang, J. Wu, and A. Shakouri, “Direct Observation of Nanoscale Peltier and Joule E ff ects at Metal − Insulator Domain Walls in Vanadium Dioxide Nanobeams,” *Nano Lett.*, vol. 14, pp. 2394–2400, 2014.

[52] K. Maize, S. R. Das, S. Sadeque, A. M. S. Mohammed, A. Shakouri, D. B. Janes, and M. A. Alam, “Super-Joule heating in graphene and silver nanowire network,” *Appl. Phys. Lett.*, vol. 106, no. 14, p. 143104, 2015.

[53] S. H. Shin, M. Masuduzzaman, M. A. Wahab, K. Maize, J. J. Gu, M. Si, A. Shakouri, P. D. Ye, and M. A. Alam, “Direct Observation of Self-heating in III-V Gate-all-around Nanowire MOSFETs,” in *Electron Device Meeting, IEEE International*, 2014, vol. 2706, no. 765, pp. 510–513.

[54] A. Ziabari, J.-H. H. Bahk, Y. Xuan, P. D. Ye, D. Kendig, K. Yazawa, P. G. Burke, H. Lu, A. C. Gossard, and A. Shakouri, “Sub-diffraction Limit Thermal Imaging for HEMT Devices,” *31th Annu. IEEE Therm. Meas. Model. Manag. Symp.*, no. 1, pp. 1–6, 2015.

[55] W. . Steen, “Principles of Optics M. Born and E. Wolf, 7th (expanded) edition, Cambridge University Press, Cambridge, 1999, 952pp. 37.50/US $59.95, ISBN 0-521-64222-1,” *Optics & Laser Technology*, vol. 32. p. 385, 2000.

[56] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. 2007.

[57] C. a Bouman, *Model Based Image Processing*. 2013.

[58] B. Vermeersch, J. Carrete, N. Mingo, A. Shakouri, and W. Lafayette, “Superdiffusive heat conduction in semiconductor alloys -- I. Theoretical foundations,” vol. 085202, pp. 1–17, 2014.

[59] G. E. Moore, “Progress in digital integrated electronics,” *1975 Int. Electron Devices Meet.*, vol. 21, 1975.

[60] R. H. DENNARD, F. H. GAENSSLEN, H.-N. YU, V. L. RIDEOVT, E. BASSOUS, and A. R. LEBLANC, “Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions,” *IEEE Solid-State Circuits Newsl.*, vol. 12, no. 1, 2007.

[61] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan, “Temperature-aware computer systems: Opportunities and challenges,” *IEEE Micro*, vol. 23, no. 6. pp. 52–61, 2003.

[62] A. H. Ajami, K. Banerjee, and M. Pedram, “Modeling and analysis of nonuniform substrate temperature effects on global ULSI interconnects,” *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 24, no. 6, pp. 849–860, 2005.

[63] Y. K. Cheng, “A Temperature-Aware Simulation Environment for Reliable ULSI Chip Design,” *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 19, no. 10, pp. 1211–1220, 2000.

[64] H. Yu, J. Ho, and L. He, “Simultaneous power and thermal integrity driven via stapling in 3D ICs,” in *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, 2006, pp. 802–808.

[65] B. Goplen and S. Sapatnekar, “Thermal via placement in 3D ICs,” in *Proceedings of the 2005 international symposium on physical design - ISPD ’05*, 2005, p. 167.

[66] “ANSYS® Academic Research, Release 15.0.” p. ANSYS® Academic Research, Release 15.0.

[67] Y. Zhan and S. S. Sapatnekar, “High-efficiency green function-based thermal simulation algorithms,” *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 26, no. 9, pp. 1661–1675, 2007.

[68] B. Wang and P. Mazumder, “Accelerated Chip-Level Thermal Analysis Using Multilayer Green’s Function,” *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 26, no. 2, 2007.

[69] C. A. Balanis, *Advanced Engineering Electromagnetics*, vol. 52. 1989.

[70] A. Ziabari, J. Park, E. K. Ardestani, J. Renau, S.-M. Kang, and A. Shakouri, “Power Blurring: Fast Static and Transient Thermal Analysis Method for Packaged Integrated Circuits and Power Devices,” *IEEE Trans. Very Large Scale Integr. Syst.*, pp. 1–1, 2014.

[71] K. Skadron, M. R. Stan, W. Huang, S. V. S. Velusamy, K. S. K. Sankaranarayanan, and D. Tarjan, “Temperature-aware microarchitecture,” *30th Annu. Int. Symp. Comput. Archit. 2003. Proceedings.*, 2003.

[72] A. Ziabari, Z. Bian, and A. Shakouri, “Adaptive Power Blurring Techniques to Calculate IC Temperature Profile under Large Temperature Variations,” in *International Microelectronic and Packaging Society (IMAPS) ATW on Thermal Management*, 2010, pp. 1–6.

[73] A. Ziabari, E. K. Ardestani, J. Renau, and A. Shakouri, “Fast thermal simulators for architecture level integrated circuit design,” in *Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2011 27th Annual IEEE*, 2011, pp. 70–75.

[74] J. Park, A. Shakouri, and S.-M. Kang, “Fast Thermal Analysis Of Vertically Integrated Circuits (3-D ICS) Using Power Blurring Method,” in *Interpack*, 2009, pp. 1–7.

[75] K. Yazawa and A. Shakouri, “Optimization of power and efficiency of thermoelectric devices with asymmetric thermal contacts,” *J. Appl. Phys.*, vol. 111, p. 024509, 2012.

[76] T. Clin, S. Turenne, D. Vasilevskiy, and R. A. Masut, “Numerical simulation of the thermomechanical behavior of extruded bismuth telluride alloy module,” in *Journal of Electronic Materials*, 2009, vol. 38, pp. 994–1001.

[77] J.-L. Gao, Q.-G. Du, X.-D. Zhang, and X.-Q. Jiang, “Thermal Stress Analysis and Structure Parameter Selection for a Bi2Te3-Based Thermoelectric Module,” *J. Electron. Mater.*, vol. 40, no. 5, pp. 884–888, Mar. 2011.

[78] E. Suhir and A. Shakouri, “Assembly Bonded at the Ends: Could Thinner and Longer Legs Result in a Lower Thermal Stress in a Thermoelectric Module Design?.",” *J. Appl. Mech.*, vol. 79.6, p. 061010, 2012.

[79] E. Suhir, *Structural Analysis in Microelectronics and Fiber Optics*. New York: Van-Nostrand, 1973.

[80] E. Suhir, “Interfacial thermal stresses in a bi-material assembly with a low-yield-stress bonding layer,” *Modelling and Simulation in Materials Science and Engineering*, vol. 14, no. 8. pp. 1421–1432, 2006.

[81] E. Suhir, “Interfacial Stresses in Bimetal Thermostats,” *Journal of Applied Mechanics*, vol. 56, no. 3. p. 595, 1989.

[82] E. Suhir and A. Shakouri, “Predicted Thermal Stress in a Multileg Thermoelectric Module (TEM) Design,” *J. Appl. Mech.*, vol. 80, p. 021012, 2013.

[83] I. O. Kulik, “Non-linear thermoelectricity and cooling effects in metallic constrictions,” *J. Phys. Condens. Matter*, vol. 6, no. 45, pp. 9737–9744, Nov. 1994.

[84] E. Bogachek, A. Scherbakov, and U. Landman, “Nonlinear Peltier effect and thermoconductance in nanowires,” *Phys. Rev. B*, vol. 60, no. 16, pp. 11678–11682, Oct. 1999.

[85] R. López and D. Sánchez, “Nonlinear heat transport in mesoscopic conductors: Rectification, Peltier effect, and Wiedemann-Franz law,” *Phys. Rev. B*, vol. 88, no. 4, p. 045129, Jul. 2013.

[86] J. Meair and P. Jacquod, “Scattering theory of nonlinear thermoelectricity in quantum coherent conductors.,” *J. Phys. Condens. Matter*, vol. 25, no. 8, p. 082201, Feb. 2013.

[87] L. P. Bulat, “Nonlinear Anisotropic Thermoelectric Energy Converter Based on Semiconductor Films,” in *Proc. ICT 2003*, 2003, pp. 372–375.

[88] L. P. Bulat, E. V Buzin, and U. S. Whang, “Physical processes in thermoelectric coolers,” in *Proceedings ICT 2001 20 International Conference on Thermoelectrics Cat No01TH8589*, 2001, pp. 435–438.

[89] G. Chen, “Potential-step amplified nonequilibrium thermal-electric converters,” *J. Appl. Phys.*, vol. 97, no. 8, p. 083707, 2005.

[90] M. Lundstrom, “Fundamentals of Carrier Transport, 2nd edn,” *Measurement Science and Technology*, vol. 13, no. 2. pp. 230–230, 2002.

[91] R. Banan Sadeghian, J.-H. Bahk, Z. Bian, and A. Shakouri, “Calculation of Nonlinear Thermoelectric Coefficients of InAs(1−x) Sb(x) Using Monte Carlo Method,” *J. Electron. Mater.*, vol. 41, no. 6, pp. 1370–1375, Dec. 2011.

[92] O. Muscato and V. Di Stefano, “Local equilibrium and off-equilibrium thermoelectric effects in silicon semiconductors,” *J. Appl. Phys.*, vol. 110, no. 9, pp. 0–10, 2011.

[93] I. Terasaki, R. Okazaki, and H. Ohta, “Search for non-equilibrium thermoelectrics,” *Scr. Mater.*, 2015.

[94] H. Elhadidy, J. Sikula, and J. Franc, “Symmetrical Current–Voltage Characteristic of a Metal–Semiconductor–Metal Structure of Schottky Contacts and Parameter Retrieval of a CdTe Structure,” *Semicond. Sci. Technol.*, vol. 27, no. 1, pp. 015006–015012, Jan. 2012.

[95] S. M. Sze, D. J. Coleman, and A. Loya, “Current Transport In Metal-Semiconductor-Metal (MSM) Structure,” *Solid State Electron.*, vol. 14, pp. 1209–1218, 1971.

[96] Z. Y. Zhang, C. H. Jin, X. L. Liang, Q. Chen, and L.-M. Peng, “Current-voltage characteristics and parameter retrieval of semiconducting nanowires,” *Appl. Phys. Lett.*, vol. 88, no. 7, p. 073102, 2006.

[97] Z. Zhang, K. Yao, Y. Liu, C. Jin, X. Liang, Q. Chen, and L.-M. Peng, “Quantitative Analysis of Current–Voltage Characteristics of Semiconducting Nanowires: Decoupling of Contact Effects,” *Adv. Funct. Mater.*, vol. 17, no. 14, pp. 2478–2489, Sep. 2007.

[98] L. K. Nanver and E. J. G. Goudena, “I-V Characteristics of Integrated n+ p n-Reachthrough Diodes,” *Solid State Electron.*, vol. 32, no. 8, pp. 637–645, 1989.

[99] D. K. Schroder, *Semiconductor Material and Device Characterization: Third Edition*. 2005.

[100] M. R. Pinto, S. Kent, A. Alam Muhammad, S. Clark, X. Wang, G. Klimeck, and D. Vasileska, “Padre.” Jan-2006.

[101] J. Christofferson and A. Shakouri, “Thermoreflectance based thermal microscope,” *Rev. Sci. Instrum.*, vol. 76, no. 2, p. 024903, 2005.

[102] J. Christofferson and A. Shakouri, “Thermoreflectance based thermal microscope,” *Rev. Sci. Instrum.*, vol. 76, 2005.

[103] X. Wang, Y. Ezzahri, J. Christofferson, and A. Shakouri, “Bias-dependent MOS transistor thermal resistance and non-uniform self-heating temperature,” *J. Phys. D. Appl. Phys.*, vol. 42, no. 7, p. 075101, 2009.

[104] K. Maize, A. Ziabari, W. D. French, P. Lindorfer, B. O’Connell, A. Shakouri, B. OConnell, and A. Shakouri, “Thermoreflectance CCD Imaging of Self-Heating in Power MOSFET Arrays,” *IEEE Trans. Electron Devices*, vol. 61, no. 9, pp. 3047–3053, Sep. 2014.

[105] S. M. Sze, *Physics of Semiconductors*, 3rd ed. Wiley, 2003.

[106] "Ioffe Physico-Technical Institute." [Online]. Available: http://www.ioffe.ru/SVA/NSM/Semicond/InP/index.html.

[107] G. J. Snyder, J. P. Fleurial, T. Caillat, R. Yang, and G. Chen, "Supercooling of Peltier cooler using a current pulse," *J. Appl. Phys.*, vol. 92, pp. 1564–1569, 2002.

[108] C. H. Lin, T. A. Merz, D. R. Doutt, M. J. Hetzer, J. Joh, J. A. Del Alamo, U. K. Mishra, and L. J. Brillson, "Nanoscale mapping of temperature and defect evolution inside operating AlGaN/GaN high electron mobility transistors," *Appl. Phys. Lett.*, vol. 95, 2009.

[109] W. D. Hu, X. S. Chen, Z. J. Quan, C. S. Xia, W. Lu, and P. D. Ye, "Self-heating simulation of GaN-based metal-oxide-semiconductor high-electron-mobility transistors including hot electron and quantum effects," *J. Appl. Phys.*, vol. 100, 2006.

[110] Y. Chang, Y. Zhang, Y. Zhang, and K. Y. Tong, "A thermal model for static current characteristics of AlGaNGaN high electron mobility transistors including self-heating effect," *J. Appl. Phys.*, vol. 99, 2006.

[111] W. Y. Zhang, R; Zhao, W. S.; Yin, "Investigation on thermo-mechanical responses in high power multi-finger AlGaN/GaN HEMTs," *Microelectron. Reliab.*, vol. 54, no. 3, pp. 575–581, 2014.

[112] G. P. Karman, M. W. Beijersbergen, A. van Duijl, D. Bouwmeester, and J. P. Woerdman, "Airy pattern reorganization and subwavelength structure in a focus," *Journal of the Optical Society of America A*, vol. 15. p. 884, 1998.

[113] S. H. Shin, M. A. Wahab, W. Ahn, A. Ziabari, K. Maize, A. Shakouri, and M. A. Alam, "Fundamental trade-off between short-channel control and hot carrier degradation in an extremely-thin silicon-on-insulator (ETSOI) technology," *Tech. Dig. - Int. Electron Devices Meet. IEDM*, vol. 2016-Febru, no. 765, pp. 20.3.1–20.3.4, 2016.

[114] K. Yazawaj, D. Kendig, K. Maizi, A. Shakour, M. Llc, S. Clara, and U. States, "Transient thermal characterization of HEMT devices," no. 1, 2014.

[115] S. Bhargava and E. Yablonovitch, "Lowering HAMR Near-Field Transducer Temperature via Inverse Electromagnetic Design," *IEEE Trans. Magn.*, vol. 51, no. 4, pp. 1–7, 2015.

[116] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and Robust Multi-Frame Super-Resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, 2004.

[117] C. Bouman, "MAP Image Restoration." [Online]. Available: https://engineering.purdue.edu/~bouman/grad-labs/MAP-Image-Restoration/.

[118] L. R. Harriott, "Limits of lithography," *Proc. IEEE*, vol. 89, no. 3, pp. 366–374, 2001.

[119] R. B. Wilson and D. G. Cahill, “Anisotropic failure of Fourier theory in time-domain thermoreflectance experiments.,” *Nat. Commun.*, vol. 5, p. 5075, 2014.

[120] A. J. Minnich, “Determining phonon mean free paths from observations of quasiballistic thermal transport,” *Phys. Rev. Lett.*, vol. 109, no. 20, pp. 1–5, 2012.

[121] J. Maassen and M. Lundstrom, “Steady-state heat transport: Ballistic-to-diffusive with Fourier’s law,” *J. Appl. Phys.*, vol. 117, no. 3, p. 035104, 2015.

[122] J. Maassen and M. Lundstrom, “A simple Boltzmann transport equation for ballistic to diffusive transient heat transport,” *J. Appl. Phys.*, vol. 117, no. 13, p. 135102, 2015.

[123] J. A. del Alamo, “Nanometre-scale electronics with III–V compound semiconductors,” *Nature*, vol. 479, no. 7373, pp. 317–323, 2011.

[124] A. Ziabari, J. Bahk, Z. Bian, H. Lu, B. Vermeersch, A. C. Gossard, and A. Shakouri, “Experimental Observation of Current-dependent Peltier Coefficient in Thin Film Microrefrigerators,” *Submitted*, 2016.

[125] D. G. Cahill, “Analysis of heat flow in layered structures for time-domain thermoreflectance,” *Rev. Sci. Instrum.*, vol. 75, no. 12, pp. 5119–5122, 2004.

[126] D. G. Cahill, “Thermal conductivity measurement from 30 to 750 K: The 3w method,” *Rev. Sci. Instrum.*, vol. 61, no. 2, pp. 802–808, 1990.

[127] C. Dames and G. Chen, “1w, 2w, and 3w Methods for Measurements of Thermal Properties,” *Rev. Sci. Instrum.*, vol. 76, no. 12, pp. 1–14, 2005.

[128] “http://www.engineersedge.com/properties_of_metals.htm,” *Engineering Edge*. [Online]. Available: http://www.engineersedge.com/properties_of_metals.htm.

[129] “http://accuratus.com/alumox.html,” *Accuratus*. .

[130] “New Semiconductor Materials, Characteristics and Properties,” *New Semiconductor Materials, Characteristics and Properties(St. Petersburg: Ioffe Physico-Technical Institute, 1998–2001)*. [Online]. Available: Http://www.ioffe.ru/SVA/NSM/Semicond/InP/thermal.html.

[131] K. T. Regner, D. P. Sellan, Z. Su, C. H. Amon, A. J. H. Mcgaughey, and J. A. Malen, “domain thermoreflectance,” *Nat. Commun.*, vol. 4, pp. 1640–1647, 2013.

[132] J. A. Johnson, A. A. Maznev, J. Cuffe, J. K. Eliason, A. J. Minnich, T. Kehoe, C. M. S. Torres, G. Chen, and K. A. Nelson, “Direct measurement of room-temperature nondiffusive thermal transport over micron distances in a silicon membrane,” *Phys. Rev. Lett.*, vol. 110, no. 2, pp. 1–5, 2013.

[133] W. Kim, S. L. Singer, A. Majumdar, D. Vashaee, Z. Bian, A. Shakouri, G. Zeng, J. E. Bowers, J. M. O. Zide, and A. C. Gossard, “Cross-plane lattice and electronic thermal conductivities of ErAs:InGaAs⁄InGaAlAs superlattices,” *Appl. Phys. Lett.*, vol. 88, no. 24, p. 242107, 2006.

[134] B. Vermeersch and G. De Mey, “A fixed-angle heat spreading model for dynamic thermal characterization of rear-cooled substrates,” in *Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, 2007, pp. 95–101.

[135] A. Shakouri, E. Y. Lee, D. L. Smith, V. Narayanmurti, and J. E. Bowers, “Thermoelectric Effects in Submicron Heterostructure Barriers,” *Microscale Thermophys. Eng.*, pp. 37–47, 1998.

APPENDICES

A. Individual via vs. via region

First, we demonstrate that using average thermal conductivity of each region based on the density of thermal via in that region is a reasonable approximation to decrease the complexity of computation. In order for this, we have used the geometry shown in Figure A.1. The dimensions are shown in the figure as well. Material properties are same as package shown in Figure 2.18 and Table 2.3.

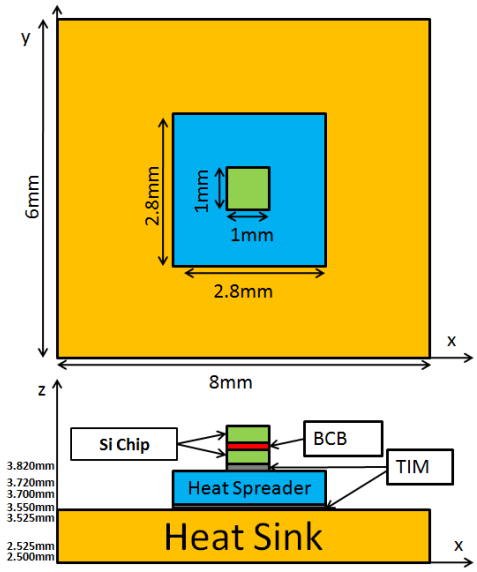


Figure A.1. Schematic of a packaged 3D IC chip.

The dies are meshed with 10µm granularity, which create 10000 surface meshed element and total of 1140000 elements for the full package. The granularity is chosen according to width of an individual via. Figure A.2a and A.2b show both a top view of the cases with individual vias in the chip and a top view of the case having thermal via regions with average thermal conductivity. The thermal conductivity of the thermal via regions (K_{via_region}) are determined based on the density of copper (Cu) vias using a weighted averaging between the thermal conductivity of Cu (K_{Cu_via}) and Si (K_{Si}) using equation (A.1).

$$K_{via\_region} = m \times K_{Cu\_via} + (1\text{-}m) \times K_{Si} \quad \text{(A.1)}$$

Here, *m* is density of copper in the region and K indicates thermal conductivity of each of the two materials (Cu and Si) indicated as subscript.

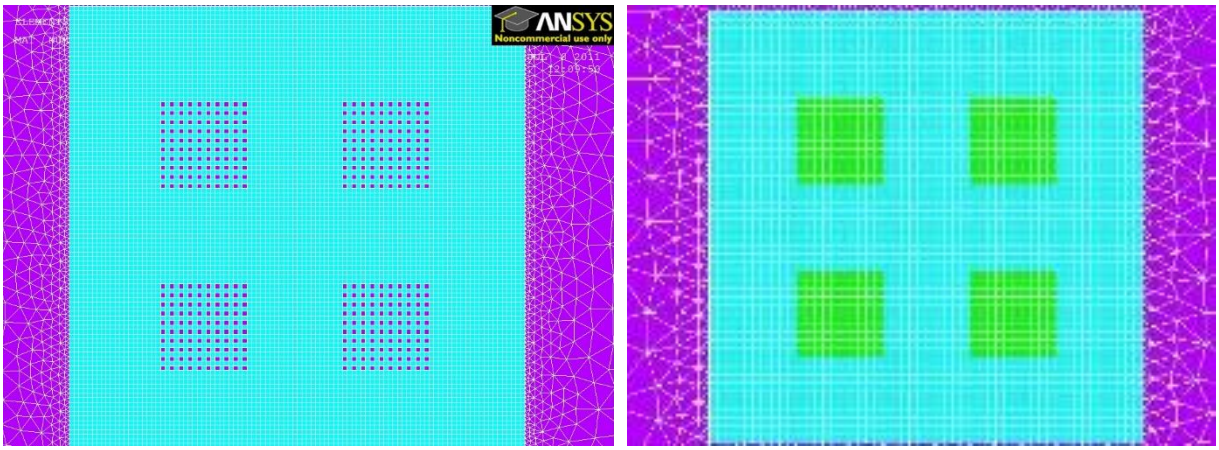


Figure A.2. A schematic from top view of the chip. (a) simulating individual vias; (b) simulating average via regions.

Two arbitrary power maps for the top and bottom layers of the 3D IC are selected, which are shown in Figure A.3. Using ANSYS, the temperature results of the top and bottom surfaces of the chip, for both configurations, are calculated. These results are illustrated in Figure A.4. As it can be seen in Figure A.4, the temperature profiles of the top layer in two cases, i.e. considering individual vias and simulating with thermal via regions, are slightly different (less than 1 °C), while those of bottom layers are almost the same.

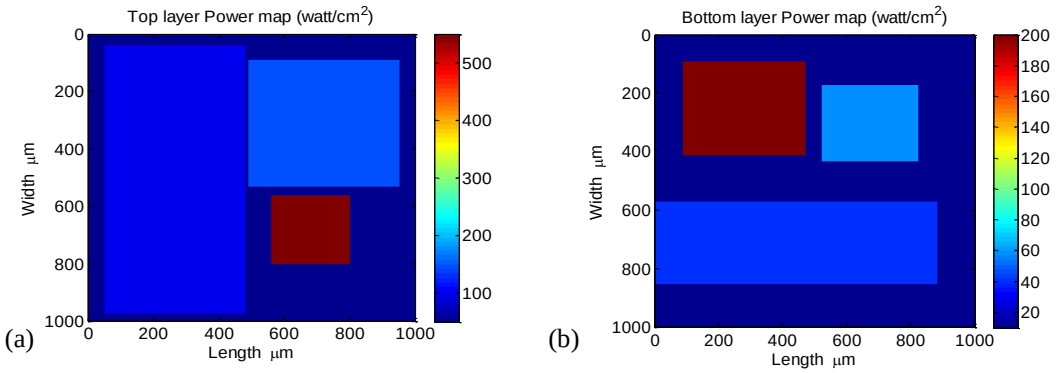


Figure A.3. Power maps in 3D IC. (a) Top active layer; (b) Bottom active layer.

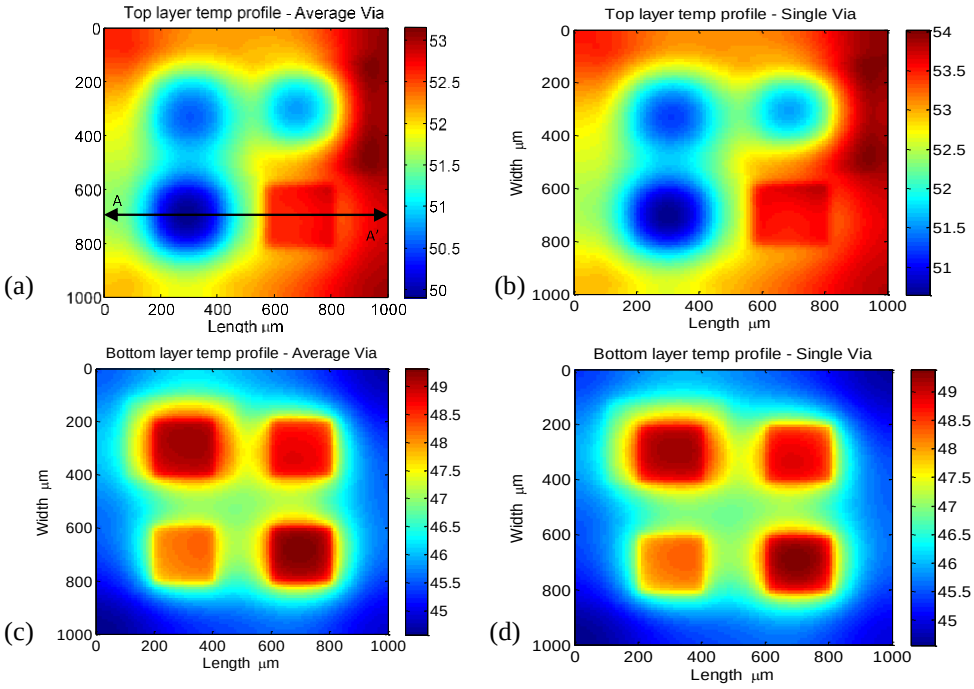


Figure A.4. Temperature profile results in different layers of the 3D IC. (a) Top layer with thermal via regions; (b) Top layer with individual vias; (c) Bottom layer with thermal via regions; (d) Bottom layer with individual vias.

The reason for the small difference might be due to the fact that heat spreading is different in the case with individual vias than the case with via region. However, the bottom line is that simulating thermal via regions instead of individual vias in the chip is a good approximation. Top and bottom temperature profiles' cross section are compared in Figure A.5. The cross sections are plotted along the path A-A', drawn in Figure A.4a.

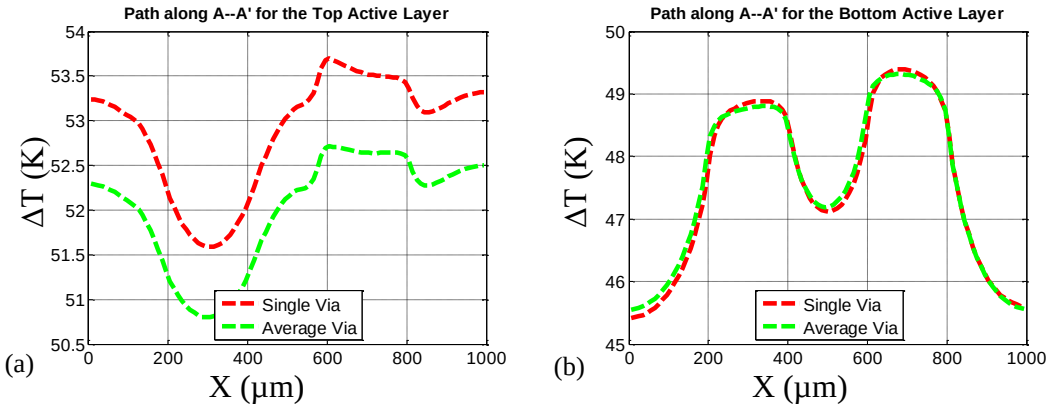


Figure A.5. Comparison between temperature profiles along path A-A' for the two cases of modeling the individual vias and thermal via regions with average thermal conductivity. Comparison at (a) Top layer, and (b) Bottom layer. Path A-A' is drawn along profile in Figure 2.11a. Results suggest that modeling either of the case will not introduce error to the modeling.

We demonstrated that calculation of weighted average of thermal conductivity is a good approximation to thermal conductivity of the thermal via regions. Thus, we can use algorithm described in Figure 2.7, to calculate thermal profile of 3D ICs including thermal vias.

B. Hybrid analytical-numerical model

A hybrid analytical-numerical scheme, based on the full heat balance equation as well as the equivalent thermal network of the microrefrigerators, is developed to separate the Peltier- and Joule -induced temperature changes, and hence extract the Peltier coefficient. A 1D thermal network with the thin-film as well as substrate thermal resistances is constructed which is shown in Figure B.1. The 3D heat spreading is included through analytical equations verified by ANSYS. We exploited this network to extract Peltier ($\Delta T_{Peltier}$) and Joule (ΔT_{Joule}) components of forward and reverse temperature changes (ΔT_{Fwd} and ΔT_{Rev}). Since measurements are done in vacuum, adiabatic boundary conditions at the surface is assumed. ΔT_{Fwd} and ΔT_{Rev} are the average temperature change on the top surface of the microcooler. In this network, R_{th_InGaAs}, R_{th_InP}, and $R_{th\_HeatSink}$ are the thermal resistances of InGaAs thin film, InP substrate, and the heat sink, respectively. Q_{J1} is the Joule heating dissipated in the InGaAs thin film and Q_{J2} is the Joule heating dissipated in the InP substrate. We calculated Q_{J1} *and* Q_{J2} from the current times voltage. This voltage includes voltages across thin film, substrate as well as the parasitic contact/interface resistances. Therefore, all the parasitics in generating Joule heating term are incorporated. Parameter α determines the percentage of the Joule heating flow back to the top surface from the InGaAs thin Film. In linear devices is α is 0.5 for both forward and reverse polarities. However, for nonlinear microrefrigerators in this work α is different for forward and reverse polarities and needs to be calculated from the following equations. We denoted them, in heat balance equations (B.1) and (B.4), as α_F and α_R for forward and reverse biasing conditions, respectively. $\Pi_1 I$ is the peltier heating at the top surface (for forward bias) and $\Pi_2 I$ is the peltier cooling at the bottom interface. Upon changing the current direction these terms change sign. Π_1 and Π_2 are effective Peltier coefficients at the top and the bottom interfaces. Different Peltier are assumed since the current densities at the two interfaces are not the same. One should note that these are the effective Peltier coefficients of the medium, which incorporate the small Seebeck from the highly doped InGaAs, and InP substrate as well as the metal in themselves. As the thermal resistance of the heat sink is very small T_3 should be equal to ambient temperature, and the Joule heating source shown in the dashed box in Figure B.1 may be neglected. This is also confirmed by the thermometer connected to the heat sink which was showing 297 about the same as the ambient temperature in the vacuum cryostat.

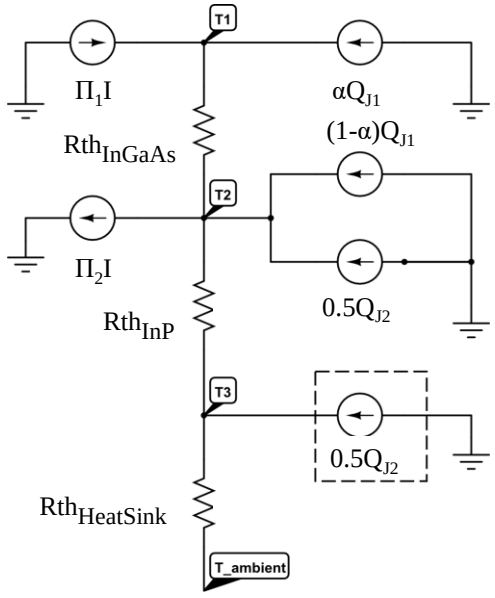


Figure B.1. Equivalent 1D thermal network of the nonlinear microrefrigerator. 3D heat spreading in the substrate and in the InGaAs thin film is included through 3D thermal resistances.

The heat balance equations for this circuit (under forward electrical biasing) are as the following:

$$\frac{T_1 - T_2}{R_{th\_InGaAs}} = \Pi_1 I + \alpha_F Q_{J1} \tag{B.1}$$

$$\frac{T_2 - T_{ambient}}{R_{th\_InP}} = -\Pi_2 I + \Pi_1 I + ((1-\alpha_F)Q_{J1} + 0.5Q_{J2}) + \alpha_F Q_{J1} \tag{B.2}$$

$$Q_{J1} = IV_{F1} \quad \& \quad Q_{J2} = IV_{F2} \tag{B.3}$$

I is the current in the device, V_{F1} is the voltage across InGaAs thin film and parasitics. V_{F2} is the voltage across the substrate. Π_1 *and* Π_2 are the effective Peltier coefficients at the top and bottom interfaces. $T_{ambient}$ is the ambient temperature. As discussed above α_F determines fraction of Joule heating dissipated in the InGaAs thin film layer that flows to the top surface.

The thermal spreading resistances, R_{th_InGaAs} and R_{th_InP}, are calculated using literature value for thermal conductivity for InP (73W/m-K) [106] and measured thermal conductivity for InGaAs layer 5W/m-K. We have used ANSYS to confirm the spreading resistances. When the device is forward biased, the top junction between the n+ cap layer and the n-type thin film is reversed biased and most of the voltage drop occurs in that junction. On the other hand, when the device is reverse biased, most of the voltage drop is on the buried n-n+ junction. Therefore, the Joule heating dissipated in R_{th_InGaAs} would be a fraction of that in the forward bias. We modified the equations for the reverse bias as follows:

$$\frac{T'_1 - T'_2}{R_{th_{InGaAs}}} = -\Pi_1 I + \alpha_R Q'_{J1} \tag{B.4}$$

$$\frac{T'_2 - T_{ambient}}{R_{th\_Inp}} = \Pi_2 I - \Pi_1 I + ((1-\alpha_R)Q'_{J1} + 0.5Q'_{J2}) + \alpha_R Q'_{J1} \tag{B.5}$$

$$Q'_{J1} = IV_{R1} \;\&\; Q'_{J2} = IV_{R2} \tag{B.6}$$

Here, α_R is the fraction of Joule heating dissipated in the thin film that flows to the top surface, V_{R1} is the voltage across thin film, and V_{R2} is the substrate voltage under reverse bias.

The following steps are taken into account to solve equations (B.1) - (B.6).

1. First finite element modeling software, ANSYS, is employed to obtain R_{th_InGaAs} and R_{th_InP} seen by heat at the top interface. T_1 and T'_1 are measured using thermoreflectance thermal imaging, and T_2 as well as T'_2 are intermediate node temperatures which will cancel out in the equations.

2. The equations 4 through 9 were solved together at each current density for different device sizes. Initially, the peltier component ($|\Pi_2 I - \Pi_1 I|$) on the buried interface was assumed zero. At low current densities the Peltier coefficient is constant and current independent, therefore we can use the nominal value of Peltier coefficient of InGaAs ($\Pi_0 = -\frac{\mu}{e} + \frac{5K_BT}{2e}$) to calculate α_F and α_R at smallest current density measurement was performed. Then, at other current densities we can use α_F and α_R and calculate the Peltier coefficients at those current densities. We obtained $\alpha_F = 0.53 \pm 0.01$ and $\alpha_R = 0.25 \pm 0.07$ for three device sizes of 75×75μm^2, 100×100 μm^2, and 120×120 μm^2.

3. Next, the Peltier component at the buried junction is taken into account. In order for this, equations (B.1) and (B.2) are combined into equation (B.7).

$$T_1 - T_{ambient} - (\alpha_F Q_{J1} R_{th_{InGaAs}} + (Q_{J1} + 0.5Q_{J2})R_{th_{InP}}) = \Pi_1 I R_{th\_InGaAs} + (-\Pi_2 I + \Pi_1 I)R_{th_{InP}} \tag{B.7}$$

In the second step above, the second term in the right hand side of the equation (10) is neglected and approximate Π_1 is calculated (let's call it Π_{eff}) Equation (B.8) was employed to distinguish between Peltier coefficient on the top and the bottom interfaces. (In implementing the procedure, we noticed that this step was not necessary and did not change the results significantly).

$$\Pi_1 = \Pi_0 + mJ_1^2,\ \Pi_2 = \Pi_0 + mJ_2^2 \tag{B.8}$$

$$J_1 = \frac{I}{A_1}, J_2 = \frac{I}{A_2} \tag{B.9}$$

Here Π_0 is the room temperature value of Peltier coefficient, m is nonlinear coefficient introduced in equation (1), and J_1 and J_2 are the current densities. A_1 is device area and A_2

is the cross sectional area seen at the buried contact. A_1/A_2 can be calculated either analytically [134], or numerically using ANSYS.

We calculated this ratio and shown it in equation (B.10):

$$\frac{A_2}{A_1} = \frac{(l+2h*tan\gamma)^2}{l^2} \tag{B.10}$$

Here l is the microrefrigerator device length, h is the thin film thickness, and γ is the spreading angle which is $\gamma \approx 45^0$ [134].

Using equations (B.7) through (B.10), we can extract the Peltier coefficient at the top interface (Π_1) as follows:

$$\Pi_1 = \Pi_{eff} + mJ_1^2\left(\frac{R_{th_{InP}}}{R_{th_{InGaAs}}}\right)\left(\frac{\left(4h^2+4hl\right)^2-2\left(4h^2+4hl\right)A_2}{A_2^2}\right) \tag{B.11}$$

Where, h is 4.45μm and l is changing from 10μm to 150μm. (Π_{eff}) is the extracted effective Peltier coefficient in step 2. Analysis shows that the second term in equation (B.11) is negligible and (Π_1) is approximately equivalent to (Π_{eff}).

4. At cryogenic temperatures, even in linear devices, α_F may be as low as 0.1 [135]. This is because of the longer mean free path of electrons due to weaker electron-phonon coupling at those temperatures. In order to estimate α_F, we assumed at small current densities (linear regime) Peltier coefficient from theory and experiment agree. From this we extracted α_F to be 0.2, 0.28 and 0.23 for 30K, 50K and 70K, respectively. Then by putting this α_F back in the equation (B.4) at larger current densities, and from the hybrid method described in steps 1-3, we calculated the Peltier coefficient at 30K, 50K, and 70K, which are plotted in Fig. 5.5b in chapter 5. We analysed the sensitivity of the extracted Peltier coefficient with respect to extracted α_F at 30K, 50K and 70K. This analysis is shown in Figure B.2.

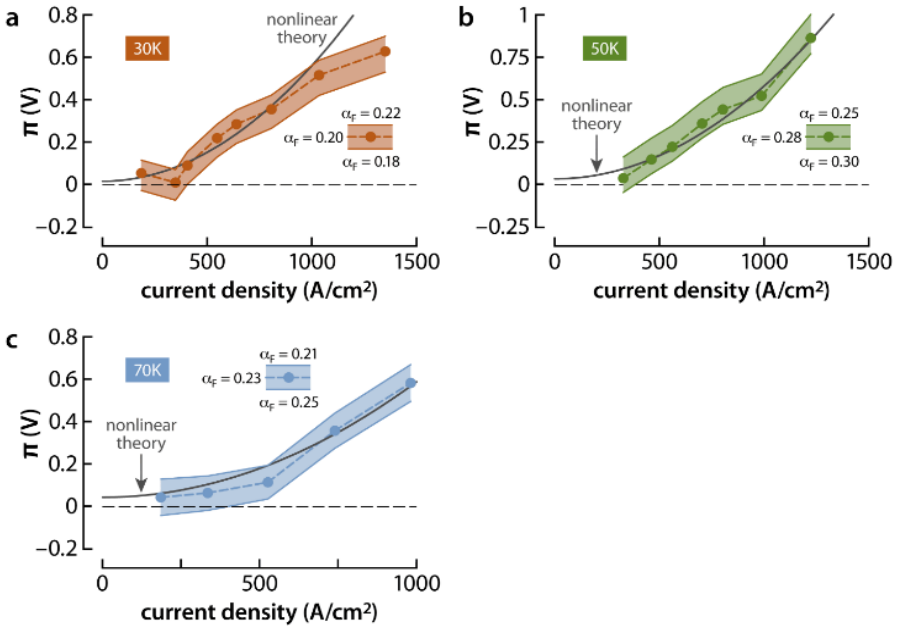


Figure B.2. Sensitivity of the extracted Peltier coefficient to the value of αF different temperatures. (a) 30K, (b) 50K, (c) 70K. The dashed line is theoretical prediction for the nonlinear Peltier coefficient.

To summarize, overall there are 9 parameters for *each device and at each current density* that we need to extract or calculate based on the measured data: Π, α_F, α_R, Q_{J1}, Q_{J2}, Q'_{J1}, Q'_{J2}, R_{th_InGaAs} and R_{th_InP}. The last two parameters obtained from experimental thermal conductivity and full 3D FEM or analytical thermal modelling of the devices. Q_{J1}, Q_{J2}, Q'_{J1}, Q'_{J2}, computed from electrical measurement (Equations B.3 and B.6). Therefore, we have three parameters Π, α_F, α_R, and 4 equations (S1, S2, S4 and S5) left. However, after simplifications these 4 equations reduces to 2 equations. Three parameters and 2 equations are left. One extra equation is needed. At low current densities nonlinear dependence of Peltier coefficient on current InGaAs is negligible, and thus we can write $\Pi = -\frac{\mu}{e} + \frac{5K_BT}{2e}$. Therefore, Π at low current densities is calculated and the 2 remaining equations are sufficient to extract α_F, and α_R easily. Then assuming these values do not depend on current densities, we extract current-dependent Peltier coefficient as a function of current. To assure that the results are accurate, the peltier coefficients obtained in forward direction (using α_F in Equations B.1 and B.2), must match the ones obtained from reverse direction (using α_R in Equations B.4 and B.5). From these the uncertainty in α_F, and α_R at different current levels were also obtained. Additionally, independent extraction of the Peltier coefficients for three different device sizes demonstrate excellent agreement and further confirms the consistency of the methodology. The same procedure is performed at cryogenic temperatures.

C. 3ω and TDTR Techniques

The principles of 3ω and TDTR techniques had already described in several references that are also referred to in the main text of this paper. Here we just present the results obtained from these techniques. Figure C.1a and b shows the results obtained from 3ω and fitting model. The extracted parameters for thermal conductivity of InGaAs, InP are 5.35W/m-K and 69.85W/m-K. The effective interface thermal conductivity of 20nm Al_2O_3 is 0.63W/m-K. InGaAs thermal conductivity and effective interface thermal conductivity are also obtained using TDTR to be 5.4±0.4, and 0.65±0.07, respectively.

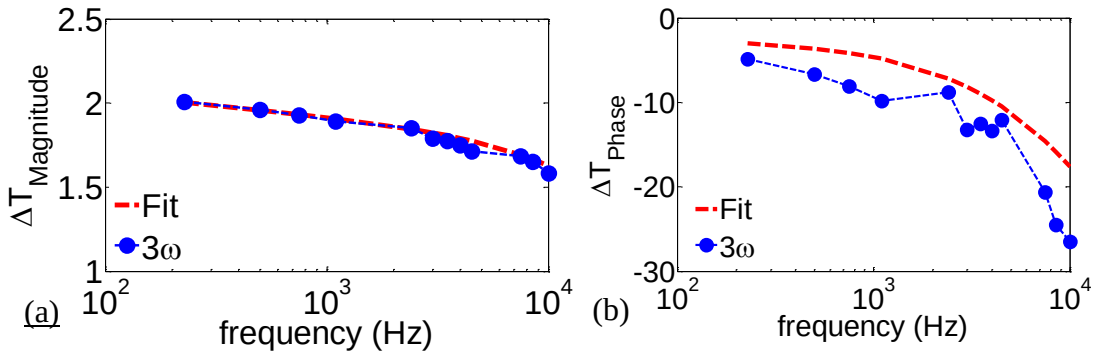


Figure C.1. 3ω Magnitude and Phase for thermal conductivity extraction. The fitted line suggests a thermal conductivity of 5.35W/m-K and 69.85W/m-K for InGaAs, InP, respectively. The effective interface thermal conductivity of 20nm Al_2O_3 is 0.63W/m-K. **a.** Magnitude; **b.** Phase.

D. Analytical Model for Nano-heater Lines on InGaAs

An analytical model based on the FEM results was developed. This is shown in Figure D.1a. FEM is done for all devices from 10µm-100nm. Knowing the amount of input heat and from the temperature change in cross plane direction, and the temperature drop across each layer, the thermal resistance of each layer ($R_{layer} = \Delta T_{layer}/Q_{in}$) can be extracted. Then by changing the thermal conductivity of each layer, we can model each layer based on their thermal conductivities. The resistance of each layer is proportional to the thermal conductivity of that layer. This analytical model was used to fit the experimental data within %0.5 and to obtain the appropriate InGaAs thermal conductivity for different size devices. The conductivity then input to full 3D FEM model with more than a million elements to obtain the temperature profile of the heater line and compare them with experiment.

It is also interesting to note how the thermal resistance of each layer contributes to the total thermal resistance. This is shown in Figure D.1b. As the size of heater line reduces the interface thermal resistance contribution become more prominent. Therefore, it was essential to assure the effective interface thermal conductivity is accurate.

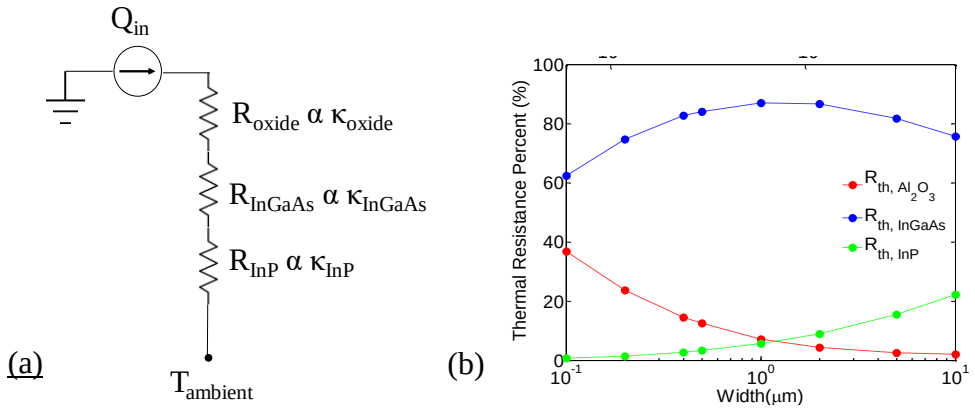


Figure D.1. Simple analytical model for thermal conductivity estimation. The results from this model is then input to full 3D FEM model to obtain the temperature profile of heater lines. **a.** Equivalent thermal network for analytical modeling. Thermal resistances of different layers are obtained knowing heat flux and junction temperature of each layer in FEM model. **b.** Contribution of thermal resistance of each layer to the total thermal resistance at different widths.

E. Extracted Gold properties

The measured properties for gold are shown in Figure E.1. Gold's electrical resistivity, thermal conductivity, coefficient of temperature dependence of resistivity, and coefficient of thermoreflectance, are plotted against device widths.

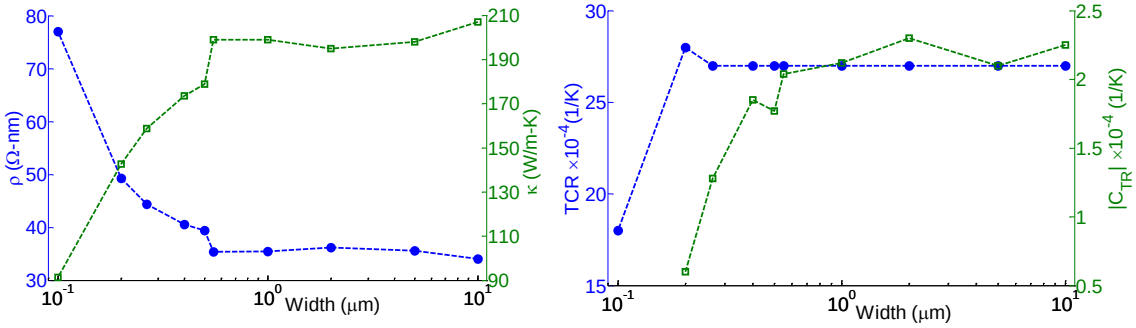


Figure E.1. Gold measured material properties. a. Electrical resistivity (ρ) and thermal conductivity (κ) of gold measured for different widths. b. Coefficient of temperature dependence of resistivity (TCR), and coefficient of thermoreflectance (C_{TR}) measured for gold at different heater line widths. For 100nm, we did not perform thermal imaging and only relied on temperature measurements based on IVT results.

VITA

VITA

Amirkoushyar Ziabari received his B.Sc and M.Sc degree in Electrical Engineering from Amirkabir University (Tehran Polytechnic), Sharif University of Technology, Tehran, Iran. In 2009 he joined University of California Santa Cruz (UCSC), Santa Cruz, California, as a graduate research assistant. During his master studies at Santa Cruz, Amirkoushyar developed a fast static and transient thermal modeling technique, Adaptive Power Blurring, for VLSI ICs.

In 2012, he received his second master's degree from UCSC and joined Purdue University, West Lafayette, IN, as a PhD student. His research interests include nanoscale electrothermal transport, applied image processing, and thermoelectric device modeling and characterization. He has developed numerical and analytical models to study micro to nanoscale electrothermal transport, and an image reconstruction scheme to reconstruct accurate temperature map of submicron size devices from their measured sub-diffraction thermal images.

LIST OF PUBLICATIONS

LIST OF PUBLICATIONS

[P1] **A. Ziabari**, J-H. Bahk, Zh. Bian, H. Lu, B. Vermeersch, A. C. Gossard, A. Shakouri, "Experimental Observation of Current-dependent Peltier Coefficient," *Phys. Rev. Lett. (Under Rev.),* 2016.

[P2] **A. Ziabari**, M. Zebarjadi, D. Vashaee, and A. Shakouri, "Nanoscale solid-state cooling: a review," *Reports Prog. Phys.*, In Press, 2016.

[P3] S. Jin, **A. Ziabari**, Y. R. Koh, M. Saei, X. Wang, B. Deng, Y. Hu, J.-H. Bahk, A. Shakouri, and G. J. Cheng, "Enhanced thermoelectric performance of P-type nanowires with pulsed laser assisted electrochemical deposition," *Extrem. Mech. Lett.*, 2016.

[P4] K. Yazawa, D. Kendig, A. Shakouri, **A. Ziabari**, A. Shakouri, "Thermal Imaging of Nanometer Features", in *InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITHERM*, 2016.

[P5] A. Shakouri, **A. Ziabari**, D. Kendig, J. Bahk, Y. Xuan, P. D. Ye, K. Yazawa, and A. Shakouri, "Stable thermoreflectance thermal imaging microscopy with piezoelectric position control," in *2016 32nd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2016, pp. 128–132.

[P6] A. Shakouri, A. S. Mohammed, Y. Koh, **A. Ziabari**, J.-H. Bahk, and B. Vermeersch, "Fractional Diffusion For Thermal Transport In Submicron Semiconductors," in *ICHMT DIGITAL LIBRARY ONLINE*, 2015.

[P7] H. Pajouhi, A. Y. Jou, R. Jain, **A. Ziabari**, A. Shakouri, C. A. Savran, and S. Mohammadi, "Flexible complementary metal oxide semiconductor microelectrode arrays with applications in single cell characterization," *Appl. Phys. Lett.*, vol. 107, no. 20, p. 203103, 2015.

[P8] S. H. Shin, M. A. Wahab, W. Ahn, **A. Ziabari**, K. Maize, A. Shakouri, and M. A. Alam, "Fundamental trade-off between short-channel control and hot carrier degradation in an extremely-thin silicon-on-insulator (ETSOI) technology," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 20–23.

[P9] A. A. Shakouri, M. E. S. Kayed, **A. Ziabari**, D. Kendig, B. Vermeersch, J. Bahk, A. A. Shakouri, M. El. Kayed, **A. Ziabari**, D. Kendig, B. Vermeersch, J. Bahk, and A. A. Shakouri, "Temperature Sensitivity and Noise in Thermoreflectance Thermal Imaging," in *31th Annual IEEE Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2015.

[P10] **A. Ziabari**, J.-H. H. Bahk, Y. Xuan, P. D. Ye, D. Kendig, K. Yazawa, P. G. Burke, H. Lu, A. C. Gossard, and A. Shakouri, "Sub-diffraction Limit Thermal Imaging for HEMT Devices," *31th Annu. IEEE Semi-Therm. Meas. Model. Manag. Symp.*, no. 1, pp. 1–6, 2015.

[P11] **A. Ziabari**, K. Yazawa, and A. Shakouri, "Designing a Mechanically Robust Thermoelectric Module for High Temperature Application," in *32nd International Thermal Conductivity Conference/ 20th International Thermal Expansion Symposium*, 2015, pp. 1–7.

[P12] T. Favaloro, **A. Ziabari**, J. H. Bahk, P. Burke, H. Lu, J. Bowers, A. Gossard, Z. Bian, and A. Shakouri, "High temperature thermoreflectance imaging and transient Harman characterization of thermoelectric energy conversion devices," *J. Appl. Phys.*, vol. 116, no. 3, 2014. 109.

[P13] **A. Ziabari**, E. Suhir, and A. Shakouri, "Minimizing thermally induced interfacial shearing stress in a thermoelectric module with low fractional area coverage," *Microelectronics J.*, vol. 45, pp. 547–553, 2014.

[P14] K. Maize, **A. Ziabari**, W. D. French, P. Lindorfer, B. O'Connell, A. Shakouri, B. OConnell, and A. Shakouri, "Thermoreflectance CCD Imaging of Self-Heating in Power MOSFET Arrays," *IEEE Trans. Electron Devices*, vol. 61, no. 9, pp. 3047–3053, Sep. 2014.

[P15] **A. Ziabari**, J.-H. Bahk, H. Lu, Z. Bian, A. C. Gossard, and A. Shakouri, "Observation of Nonlinear Peltier Cooling in Thermoelectric Microrefrigerators," in *Materials Research Society Symposium (MRS), Spring*, 2014.

[P16] **A. Ziabari**, J.-H. Bahk, H. Lu, Z. Bian, A. C. Gossard, and A. Shakouri, "Observation of Nonlinear Peltier Coefficient in Low-doped n-type InGaAs at Cryogenic Temperatures," in *International Conference on Thermoelectrics, ICT*, 2014.

[P17] **A. Ziabari**, J. Park, E. K. Ardestani, J. Renau, S.-M. Kang, and A. Shakouri, "Power Blurring: Fast Static and Transient Thermal Analysis Method for Packaged Integrated Circuits and Power Devices," *IEEE Trans. Very Large Scale Integr. Syst.*, pp. 1–1, 2014.

[P18] E. K. Ardestani, **A. Ziabari**, A. Shakouri, and J. Renau, "Enabling power density and thermal-aware floorplanning," in *28th Annual IEEE Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2012, pp. 302–307.

[P19] **A. Ziabari** and A. Shakouri, "Fast thermal simulations of vertically integrated circuits (3D ICs) including thermal vias," in *InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITHERM*, 2012, pp. 588–596.

[P20] **A. Ziabari**, E. Suhir, and A. Shakouri, "Minimizing thermally induced interfacial shearing stress in a thermoelectric module," in *18th International workshop on thermal investigations of ICS and systems (Therminic)*, 2012.

[P21] K. Yazawa, **A. Ziabari**, A. Shakouri, V. Sahu, A. G. Fedorov, and Y. Joshi, "Cooling power optimization for hybrid solid-state and liquid cooling in integrated circuit chips with hotspots," *13th Intersoc. Conf. Therm. Thermomechanical Phenom. Electron. Syst.*, pp. 99–106, May 2012.

[P22] V. Sahu, A. G. Fedorov, Y. Joshi, K. Yazawa, **A. Ziabari**, and A. Shakouri, "Energy efficient liquid-thermoelectric hybrid cooling for hot-spot removal," in *28th Annual IEEE Thermal Measurement, Modeling & Management Symposium (SEMI-THERM)*, 2012, pp. 130–134.

[P23] **A. Ziabari**, E. K. Ardestani, J. Renau, and A. Shakouri, "Fast thermal simulators for architecture level integrated circuit design," in *Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2011 27th Annual IEEE*, 2011, pp. 70–75.

[P24] **A. Ziabari**, Z. Bian, and A. Shakouri, "Adaptive Power Blurring Techniques to Calculate IC Temperature Profile under Large Temperature Variations," in *International Microelectronic and Packaging Society (IMAPS) ATW on Thermal Management*, 2010, pp. 1–6.